MST 204.20

The Open University
Mathematics/Science/Technology
An Inter-faculty Second Level Course
MST204 Mathematical Models and Methods

mathematical models and methods

unit 20

Matrix algebra and determinants

The Open University

Mathematics/Science/Technology
An Inter-faculty Second Level Course
MST204 Mathematical Models and Methods

Unit 20
Matrix algebra and determinants

Prepared for the Course Team
by Jen Phillips

The Open University

The Open University, Walton Hall, Milton Keynes.

First published 1981. Reprinted 1983, 1985, 1989, 1992, 1994, 1996.

ISBN 0 335 14049 1

Printed and bound in the United Kingdom by Staples Printers Rochester Limited, Neptune Close, Medway City Estate, Frindsbury, Rochester, Kent ME2 4LT.

This text forms part of the correspondence element of an Open University Second Level Course.

For general availability of supporting material referred to in this text, please write to: Open University Educational Enterprises Limited, 12 Cofferidge Close, Stony Stratford, Milton Keynes, MK11 1BY, Great Britain.

Further information on Open University courses may be obtained from The Admissions Office, The Open University, P.O. Box 48, Milton Keynes, MK7 6AB.

1.7

Contents

Introduction

When we were solving sets of simultaneous equations like

$$\begin{aligned} 2x_1 + 3x_2 - 4x_3 &= 5 \\ x_1 + 2x_2 + x_3 &= 3 \\ 3x_1 - x_2 + 4x_3 &= 6, \end{aligned}$$

we found we were particularly interested in the rectangular blocks of numbers

$$\begin{bmatrix} 2 & 3 & -4 \\ 1 & 2 & 1 \\ 3 & -1 & 4 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 5 \\ 3 \\ 6 \end{bmatrix}$$

These rectangular blocks are called **matrices** (plural of **matrix**). There are other situations where such blocks of numbers occur. For instance, look at the information in Table 1 below showing the prices of different articles in competing shops. (Prices are given in pence.)

Table 1

	Knife	Fork	Dessert spoon	Dessert fork	Small knife
Shop A	60	52	48	48	58
Shop B	65	53	45	45	57
Shop C	63	54	47	49	56

The numerical information in this table is contained in the matrix

$$\begin{bmatrix} 60 & 52 & 48 & 48 & 58 \\ 65 & 53 & 45 & 45 & 57 \\ 63 & 54 & 47 & 49 & 56 \end{bmatrix}$$

You will meet a matrix like this in Section 2.

Another example of a simple matrix can be obtained from the vector

$$\mathbf{r} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}.$$

The essential information about the magnitude and direction of $\mathbf{r}$ is contained in the three numbers a_1, a_2 and a_3. So it is quite usual to represent $\mathbf{r}$ by a **column matrix** $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$, or sometimes by a **row matrix** $[a_1 \quad a_2 \quad a_3]$.

It is because vectors are often represented by such column (or row) matrices, that such matrices are in fact usually referred to as **column vectors** (or **row vectors**). These terms are extended to refer to any matrix consisting of a single column (or row) of any length. For example, $[1 \quad 2\frac{1}{2} \quad 4 \quad 7]$ is referred to as a row vector.

Manipulation of matrices like the ones above occurs sufficiently often for it to be worth while to develop an *algebra* of matrices which will enable us to refer simply and briefly to the manipulations we carry out.

You saw in the unit on complex numbers how it was possible to create an algebra for things other than real numbers, with suitable definitions for '+' and '×'. We can manipulate complex numbers in the same way as real numbers, simply because they obey the same laws. When we have defined what we mean by matrix 'addition' and 'multiplication', we will find that the rules for matrix algebra are largely similar to those for real and complex numbers. The main purpose of this unit is to investigate these and to learn how to use the resulting algebra. For comparison, the main laws for real numbers are given in the *Handbook*.

Study guide

You must make sure that you understand the contents of Sections 1 and 2 and Subsection 3.1 (the tape subsection) before you watch the television programme. This is probably about three hours' work.

If you have studied matrices before, you should find this a reasonably easy unit. It is for this reason that there are not too many matrix multiplication examples included in the body of the text. However, if you have never met matrix multiplication before, you will probably find that you need to do some more examples, as it takes a bit of time to get facility with this operation. It is for this reason that there are plenty of multiplication examples at the end of Section 2. In general, the end of section exercises should be used, either as extra practice if you need it, or for revision purposes. The end of unit test is there for you to see if you have assimilated the contents of the unit!

It is expected that you will study Sections 1 and 2 in the first working session. This will be a bit heavy if you haven't studied matrices before, but some of the later sections will take much less time.

Finally, Section 5 is about **determinants**. Determinants are *not* matrices, but to study this section you need to know the definition of a matrix.

1 Some simple matrix operations

1.1 Matrix notation

You will remember from *Unit 9* on simultaneous equations that a **matrix** is a rectangular block of numbers. When we wanted to refer to a general set of two equations in two unknowns, we wrote

$$a_1x_1 + b_1x_2 = c_1$$
$$a_2x_1 + b_2x_2 = c_2,$$

and when we wished to use the matrix of the left-hand side coefficients we wrote

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}.$$

This notation is fine when discussing small sets of equations or small matrices, but we would run out of letters of the alphabet if we wished to talk about larger matrices. It is for this reason that we adopt the notation

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2.$$

This very convenient notation means that we never run out of names for the coefficients, and can cope with any number of equations or any size matrix in this way.

It is usual practice to use a capital letter when referring to matrices. In this course, we shall use a capital letter in bold type. Thus the matrix

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

will be referred to briefly as **A**. Other matrices could be called **X** or **B**, for example. The only exception to this notation is that used for matrices consisting of a single column. It is common practice to refer to these 'column vectors' using a lower-case letter. We shall use a lower-case letter in bold type for these. Thus the column vector

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

will be referred to briefly as **x**.

Now, when we use the letter **A** to represent a matrix, it tells us nothing about its size or its contents. Suppose we wish to refer to a matrix which has two rows and three columns. We call this a *two by three matrix* and write it as '2×3'.

Note: it is important to specify the number of rows before the number of columns.

Exercise 1

Write down examples of matrices of the following size.

(i) 2×2 (ii) 3×4 (iii) 4×1 (iv) 1×2.

[*Solution on p. 55*]

Each individual entry in a matrix is called an **element**. The element q_{ij} in a matrix **Q** refers to the element in the ith row and jth column.

Example 1

The elements of a 2×4 matrix **Q** are labelled in the following way:

$$\mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{21} & q_{22} & q_{23} & q_{24} \end{bmatrix}.$$

In this way, each individual element in a given matrix has its own name, and we know where to find it.

Exercise 2

How would we refer to the element in the fourth row and fifth column of a matrix **P**?

[*Solution on p. 55*]

In general, a matrix **A** with m rows and n columns is called an $\boldsymbol{m \times n}$ **matrix**:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

In this matrix, a_{ij} is the element in the ith row and jth column.

In the case of a column vector **x**, it is only necessary to specify one subscript. For instance, we know that x_3 is the element in the third row of a column vector **x**.

1.2 Equality of matrices

We define '$\mathbf{A} = \mathbf{B}$' to mean that **A** and **B** are *identical*, i.e.

(a) that the matrices **A** and **B** are the same size,

(b) that the corresponding elements are equal.

So, if $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 4 & -1 \end{bmatrix}$,

then $\mathbf{A} = \mathbf{B}$ means that

$$\begin{aligned} a_{11} &= 1, & a_{12} &= 2, & a_{13} &= 1, \\ a_{21} &= 0, & a_{22} &= 4, & a_{23} &= -1. \end{aligned}$$

In this instance, the single matrix equation provides us with a shorthand notation for writing down six real-number equations.

Exercise 3

Write down the real-number equations implied by the statement $\mathbf{A} = \mathbf{B}$, if:

(i) $\mathbf{A} = \begin{bmatrix} r & u \\ z & w \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 4 & 9 \\ -3 & 4 \end{bmatrix}$.

(ii) $\mathbf{A} = \begin{bmatrix} ax + by \\ dx + ey \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} c \\ f \end{bmatrix}$.

[*Solution on p. 55*]

1.3 Matrix addition

You know from *Unit 14* on vector algebra, that the sum of the two vectors

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \quad \text{and} \quad \mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$$

is

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1)\mathbf{i} + (a_2 + b_2)\mathbf{j} + (a_3 + b_3)\mathbf{k}.$$

In our matrix notation, this would be written as

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{bmatrix}.$$

So, if this matrix representation of vectors is to be at all useful, the definition of matrix addition would have to include this statement above. A fairly intuitive extension of this gives us the following definition.

> If $\mathbf{A}$ and $\mathbf{B}$ are *two matrices of the same size*, we define
>
> $$\mathbf{A} + \mathbf{B}$$
>
> to mean the single matrix formed by adding the corresponding elements of $\mathbf{A}$ and $\mathbf{B}$.

Example 2

$$\begin{bmatrix} 2 & 3 \\ 1 & 4 \\ -1 & 5 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 4 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 2+1 & 3+0 \\ 1+0 & 4+4 \\ -1+1 & 5+3 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 1 & 8 \\ 0 & 8 \end{bmatrix}.$$

However, if

$$\mathbf{A} = [1 \quad 2] \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},$$

we would not be able to add these matrices as, by definition, matrices *must* be the same size before we can add them.

Another example of matrix addition was seen in the Gaussian elimination method in *Unit 9*. Each row of the matrix we used in the elimination process can itself be considered as a matrix. We were quite cheerfully adding one row to another so as to obtain a new row during each stage of the elimination process.

Exercise 4

Given that

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 3 \\ 4 & 3 & 2 \\ 1 & 5 & 6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -1 & 4 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 2 & 3 \\ 1 & 4 \\ 6 & 3 \end{bmatrix},$$

$$\mathbf{x} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} a \\ b \end{bmatrix},$$

find (where possible):

(i) $\mathbf{A} + \mathbf{B}$, (ii) $\mathbf{A} + \mathbf{C}$, (iii) $\mathbf{C} + \mathbf{x}$, (iv) $(\mathbf{x} + \mathbf{y}) + \mathbf{z}$.

Exercise 5

Write down the real-number equations implied by $\mathbf{x} + \mathbf{y} = \mathbf{z}$, where $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ are the matrices defined in Exercise 4.

Exercise 6

A matrix $\mathbf{C}$ is formed by adding two $m \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$.

(i) What size will $\mathbf{C}$ be?

(ii) How do we form the element c_{ij}?

[*Solutions to Exercises 4–6 on p. 55*]

Before we can do any algebra with our matrices, we must check whether the rules for addition which operate for real numbers are true for matrices. In the following proofs we shall assume that the matrices are of the same size.

1. *To prove* the commutative law for matrices for addition:

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}.$$

Proof

A typical element in the ith row and jth column of $\mathbf{A} + \mathbf{B}$ is $a_{ij} + b_{ij}$. The equivalent element for $\mathbf{B} + \mathbf{A}$ is $b_{ij} + a_{ij}$.

So, as we know from the commutative law for real numbers that $a_{ij} + b_{ij} = b_{ij} + a_{ij}$, it follows that

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}.$$

2. *To prove* the associative law for matrices for addition:

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}.$$

Proof

A 'typical' element of $\mathbf{A} + (\mathbf{B} + \mathbf{C})$ is $a_{ij} + (b_{ij} + c_{ij})$. The equivalent element in $(\mathbf{A} + \mathbf{B}) + \mathbf{C}$ is $(a_{ij} + b_{ij}) + c_{ij}$. Now, from the associative law for real numbers, we know that

$$a_{ij} + (b_{ij} + c_{ij}) = (a_{ij} + b_{ij}) + c_{ij}.$$

Hence we derive the associative rule for matrices:

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}.$$

Having proved these two laws for matrices, we can now use them when required.

It is useful at this stage to define the **zero matrix 0**, and the **negative** of a matrix $\mathbf{A}$. These are analogous to the number zero, and the negative of a number a, in our ordinary number system.

> We define a **zero matrix** as a matrix consisting entirely of zeros. Such a matrix (whatever its size) is denoted by $\mathbf{0}$, and has the property that
>
> $$\mathbf{A} + \mathbf{0} = \mathbf{A}.$$

The size of a zero matrix is fixed by the context in which it arises. For instance, if $\mathbf{A}$ above were an $m \times n$ matrix, then the '$\mathbf{0}$' would also be an $m \times n$ matrix.

> A matrix $-\mathbf{A}$ is one whose elements are the negatives of those of $\mathbf{A}$. It has the property that
>
> $$\mathbf{A} + (-\mathbf{A}) = \mathbf{0}.$$

For instance, if $\mathbf{A} = [1 \quad -2]$, then $-\mathbf{A} = [-1 \quad 2]$.

Whenever we write $\mathbf{A} - \mathbf{B}$, this will be taken to mean $\mathbf{A} + (-\mathbf{B})$. In practice we just subtract the corresponding elements of $\mathbf{B}$ from $\mathbf{A}$.

Exercise 7

Given that $\mathbf{A} = \begin{bmatrix} -2 & -1 \\ -4 & 1 \end{bmatrix}$, write down the matrix $-\mathbf{A}$, and the zero matrix such that $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$.

Exercise 8

If $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 1 & 2 \\ 5 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 0 & 3 \\ -1 & 1 \\ 2 & -3 \end{bmatrix}$, find $\mathbf{A} - \mathbf{B}$.

[*Solutions to Exercises 7 and 8 on p. 55*]

1.4 Multiplication of a matrix by a number

The effect of performing the operation

$$\mathbf{A} + \mathbf{A}$$

is to obtain a matrix whose elements are double those of **A**. Similarly, if we add $\mathbf{A} + \mathbf{A} + \mathbf{A}$, we obtain a matrix whose elements are three times those of **A**. It is tempting to think of $\mathbf{A} + \mathbf{A}$ as $2\mathbf{A}$, and $\mathbf{A} + \mathbf{A} + \mathbf{A}$ as $3\mathbf{A}$, and we extend this idea to make the following definition.

> If k is a real number, and **A** is any matrix, then we define
>
> $k\mathbf{A}$ (or $\mathbf{A}k$)
>
> to be a matrix (which is the same size as **A**) in which each element is the corresponding element of **A** multiplied by k.

It should be noted from this definition that

$$(-1)\mathbf{A} = -\mathbf{A} \qquad \text{and} \qquad 0\mathbf{A} = \mathbf{0}.$$

Example 3

If $\mathbf{A} = \begin{bmatrix} 2 & 1 & 3 \\ 4 & 2 & 0 \end{bmatrix}$, then $5\mathbf{A} = \begin{bmatrix} 10 & 5 & 15 \\ 20 & 10 & 0 \end{bmatrix}$.

With this definition of multiplication of a matrix by a number, it can be shown that:

(i) $(k_1k_2)\mathbf{A} = k_1(k_2\mathbf{A})$,

(ii) $k(\mathbf{A} + \mathbf{B}) = k\mathbf{A} + k\mathbf{B}$,

(iii) $(k_1 + k_2)\mathbf{A} = k_1\mathbf{A} + k_2\mathbf{A}$.

These laws, together with the laws for matrix addition, mean that, provided our matrices are of the same size and we confine ourselves to the algebra in this section, we can manipulate matrix algebra in much the same way as real number algebra.

Exercise 9

Given $k = 5$, $m = 2$, $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 3 & 2 \\ 4 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 1 & -1 \end{bmatrix}$, write as a single matrix the following expressions:

(i) $k\mathbf{A}$,

(ii) $k(\mathbf{A} + m\mathbf{B})$,

(iii) $k\mathbf{A} - m\mathbf{B}$,

(iv) $m\mathbf{A} + m\mathbf{B} + k\mathbf{A} - k\mathbf{B}$.

[*Solution on p. 55*]

Example 4

We can find the matrix **X** from the matrix equation

$$2(\mathbf{A} + 3\mathbf{X}) - \mathbf{A} = \mathbf{B} + \mathbf{A}$$

by using the laws and rules of this section. The references are to the list of rules given in the summary of this section.

$$2\mathbf{A} + 6\mathbf{X} - \mathbf{A} = \mathbf{B} + \mathbf{A} \qquad \text{(Summary 3(iv) then (iii))}$$
$$6\mathbf{X} + 2\mathbf{A} - \mathbf{A} = \mathbf{B} + \mathbf{A} \qquad \text{(Summary 3(i))}$$
$$6\mathbf{X} + \mathbf{A} = \mathbf{B} + \mathbf{A}. \qquad \text{(Summary 3(v))}$$

Adding $-\mathbf{A}$ to both sides of the equation, we get

$$6\mathbf{X} = \mathbf{B} \qquad \text{(Summary 3(v))}$$

and dividing both sides by 6, we get

$$\mathbf{X} = \tfrac{1}{6}\mathbf{B}. \qquad \text{(Summary 3(iii))}$$

In fact, as you can see, this solution involves only the operations discussed in this section. We work with this equation in exactly the same way as we would with an equation involving numbers.

Exercise 10

(i) Simplify the expression

$$2(\mathbf{A} + 3\mathbf{B}) - 4(\mathbf{A} - \mathbf{B}).$$

(ii) Solve the following equation for $\mathbf{X}$:

$$\mathbf{X} + 3(\mathbf{A} + \mathbf{X}) = 2\mathbf{X} - \mathbf{A}.$$

[*Solution on p. 55*]

Summary of Section 1

1. Notation

(i) A **matrix** is represented by a capital letter in bold type, e.g. $\mathbf{A}$, $\mathbf{B}$, ...

(ii) **Column vectors** are represented by lower case letters in bold type, e.g. $\mathbf{a}, \mathbf{b}, \ldots$

(iii) A matrix $\mathbf{A}$ with m rows and n columns is called an $\boldsymbol{m \times n}$ **matrix**.

(iv) a_{ij} is the element in the ith row and jth column of a matrix $\mathbf{A}$.

2. Algebra

(i) Equality: $\mathbf{A} = \mathbf{B}$ means that $\mathbf{A}$ and $\mathbf{B}$ are the same size, and their corresponding elements are equal.

(ii) Matrix addition: if two matrices $\mathbf{A}$ and $\mathbf{B}$ are the same size, then

$$\mathbf{A} + \mathbf{B}$$

is the single matrix formed by adding the corresponding elements of $\mathbf{A}$ and $\mathbf{B}$.

(iii) $k\mathbf{A}$ is the matrix whose elements are k times those of $\mathbf{A}$.

Two particular matrices are defined:

(iv) a matrix $\mathbf{0}$ (consisting entirely of zeros) with the property that $\mathbf{A} + \mathbf{0} = \mathbf{A}$;

(v) a matrix $-\mathbf{A}$ whose elements are the negatives of those of $\mathbf{A}$. It has the property that $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$.

3. Laws

Given matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ (all the same size), then:

(i) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$,

(ii) $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$,

(iii) $(k_1k_2)\mathbf{A} = k_1(k_2\mathbf{A})$,

(iv) $k(\mathbf{A} + \mathbf{B}) = k\mathbf{A} + k\mathbf{B}$,

(v) $(k_1 + k_2)\mathbf{A} = k_1\mathbf{A} + k_2\mathbf{A}$.

In this way, we start to build up an algebra for matrices very similar to the algebra for real numbers.

End of section exercise

Exercise 11

Let $\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 7 & -1 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

(i) Find, if possible,
(a) $\mathbf{A} + \mathbf{B}$, (b) $\mathbf{B} - \mathbf{C}$, (c) $\mathbf{C} + 2\mathbf{x}$ (d) $\mathbf{x} + 2\mathbf{y}$, (e) $\mathbf{A} - 2\mathbf{B} + \mathbf{C}$.

(ii) Find the matrix $\mathbf{P}$ if
$2(\mathbf{P} + 2\mathbf{A}) + \mathbf{B} = 3\mathbf{C}$.

[*Solution on p. 55*]

2 Matrix multiplication

2.1 Motivation

Example 1

	milk (pints)	bread (loaves)	sugar (kg)
John	6	3	2
Jane	10	2	2
Joyce	7	2	1
Jim	15	4	3

Table 1

	A	B
milk	18	17
bread	30	32
sugar	35	33

Table 2

Table 1 above shows four people's weekly milk, bread and sugar requirements, and Table 2 shows the cost (in pence) of these items in two supermarkets A and B. There is another table which can be constructed from these two tables—a table which shows how much the four people will have to pay for their requirements in either supermarket.

	A	B
John	268	
Jane		
Joyce		216
Jim		

Table 3

To work out how much it would cost John to shop in A, he buys

6 pints of milk at 18p each,
3 loaves at 30p each, and
2 kg sugar at 35p each,

giving a total cost of

$$(6 \times 18) + (3 \times 30) + (2 \times 35)$$
$$= 268\text{p}.$$

To work out how much it would cost Joyce to shop in B, we take Joyce's row in Table 1 (her requirements), and the 'B' column in Table 2 (the cost of these requirements), and obtain the cost

$$(7 \times 17) + (2 \times 32) + (1 \times 33)$$
$$= 216\text{p}.$$

Exercise 1

Fill in the rest of the entries in Table 3.

[Solution on p. 56]

This combination of the two tables could be expressed in matrix terms:

$$\begin{bmatrix} 6 & 3 & 2 \\ 10 & 2 & 2 \\ 7 & 2 & 1 \\ 15 & 4 & 3 \end{bmatrix} \text{ combined with } \begin{bmatrix} 18 & 17 \\ 30 & 32 \\ 35 & 33 \end{bmatrix} \text{ gives } \begin{bmatrix} 268 & 264 \\ 310 & 300 \\ 221 & 216 \\ 495 & 482 \end{bmatrix}.$$

This process, of combining rows of the left-hand matrix with columns of the right-hand matrix to produce elements of a new matrix, is the operation which we call **matrix multiplication**. This is probably the operation which we perform most frequently when using matrices. You will see some examples of its use in this unit.

$$\begin{bmatrix} & & \\ a_1 & a_2 & a_3 \\ & & \end{bmatrix}\begin{bmatrix} & b_1 & \\ & b_2 & \\ & b_3 & \end{bmatrix} = \begin{bmatrix} & & \\ & \square & \\ & \uparrow & \end{bmatrix}$$
$$a_1b_1 + a_2b_2 + a_3b_3$$

You will notice the analogy with the formula for calculating scalar products of vectors.

2.2 The definition of matrix multiplication

When we write

$$\mathbf{C} = \mathbf{AB},$$

we mean we are combining (multiplying) two matrices **A** and **B** to obtain a new matrix **C**. Each element of this matrix is formed by taking a row of the left-hand matrix **A**, and combining it with a column of the right-hand matrix **B**, along the lines indicated in the previous section. The following example should show how the individual elements of such a matrix **C** are formed.

Example 2

Find **AB** if

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 6 & -1 \\ 0 & 2 \\ -1 & 1 \end{bmatrix}.$$

Solution

Let

$$\mathbf{AB} = \mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix},$$

where c_{11} is formed by taking row 1 of **A** and combining it with column 1 of **B**, giving

$$c_{11} = (2 \times 6) + (1 \times 0) + (3 \times -1) = 9.$$

To get c_{12}, combine row 1 of **A** and column 2 of **B**, to give

$$c_{12} = (2 \times -1) + (1 \times 2) + (3 \times 1) = 3.$$

To get c_{21}, combine row 2 of **A** and column 1 of **B**, to give

$$c_{21} = (4 \times 6) + (5 \times 0) + (6 \times -1) = 18.$$

To get c_{22}, combine row 2 of **A** and column 2 of **B**, to give

$$c_{22} = (4 \times -1) + (5 \times 2) + (6 \times 1) = 12.$$

Thus we obtain

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 6 & -1 \\ 0 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 3 \\ 18 & 12 \end{bmatrix},$$

$$\mathbf{A} \qquad\qquad \mathbf{B} \quad = \quad \mathbf{C}.$$

So, to find an element c_{ij} of a matrix $\mathbf{C} = \mathbf{AB}$, *take row i of the left-hand matrix* **A**, *and combine it with column j of the right-hand matrix* **B** in the manner described above.

Exercise 2

Find **AB** if

$$\mathbf{A} = \begin{bmatrix} 4 & 7 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}.$$

[*Solution on p. 56*]

There are more exercises on matrix multiplication later in this section. Make certain that you do enough of them to feel confident that you can do the process.

The method above actually determines the size of $\mathbf{C} = \mathbf{AB}$. For instance in Example 2, to obtain an element c_{31}, we would have had to combine row 3 of **A** with row 1 of **B**. As **A** has no third row, neither has **C**. So we conclude that **C** *has the same number of rows as the left-hand matrix* **A**. Similarly, to obtain an element c_{13}, we would have to combine row 1 of **A** with column 3 of **B**. As **B** has no third column, neither has **C**. So we conclude that **C** *has the same number of columns as the right-hand matrix* **B**.

Example 3

If **A** is a 2×3 matrix, and **B** is a 3×4 matrix, then **AB** will be a 2×4 matrix. For instance,

$$\underset{2 \times 3}{\begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \end{bmatrix}} \underset{3 \times 4}{\begin{bmatrix} 6 & -1 & 3 & 1 \\ 0 & 2 & 1 & 1 \\ -1 & 1 & 0 & 1 \end{bmatrix}} = \underset{2 \times 4}{\begin{bmatrix} 9 & 3 & 7 & 6 \\ 18 & 12 & 17 & 15 \end{bmatrix}}.$$

Exercise 3

Find the size of **AB**, if **A** is a 3×5 matrix, and **B** is a 5×2 matrix.

[*Solution on p. 56*]

Finally, the rule for combining the elements of **A** and **B** has automatically imposed a restriction on the size of matrices we can multiply together. For instance, if a typical row of **A** were $[a_1 \quad a_2 \quad a_3 \quad a_4]$ and a typical column of **B** were $\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, then we wouldn't be able to do the specified row-column operation, as we would have an element a_4 left over. So we have the natural restriction, that *to be able to form* **AB**, *the number of columns of* **A** *must equal the number of rows of* **B**.

Example 4

If **A** is a 2×3 matrix, then if **AB** is to be formed, **B** will have to have 3 rows. For instance, the products

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 6 & -1 \\ 0 & 2 \\ -1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ -1 \end{bmatrix},$$

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 6 & -1 & 1 \\ 0 & 2 & 2 \\ -1 & 1 & 3 \end{bmatrix}, \text{ etc.}$$

can all be formed, whereas to form the product

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \end{bmatrix}\begin{array}{l}\begin{bmatrix} 6 & -1 \\ 0 & 2 \end{bmatrix}\\ \;\;?\end{array}$$

is not possible.

(Note, there is no restriction on the number of *columns* that **B** may have, or the number of *rows* that **A** may have.)

Exercise 4

If **A** is a 2×4 matrix, then what size must **B** be to be able to form

(i) **AB** (ii) **BA**?

[*Solution on p. 56*]

The restrictions on the sizes can be summed up as follows.

$$\begin{array}{ccc} \mathbf{A} & \mathbf{B} & = & \mathbf{C} \\ m \times \boxed{n} & \boxed{n} \times p & & m \times p \end{array}$$

Exercise 5

Given the matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 2 \\ 0 \end{bmatrix},$$

(i) state which matrix products can be formed,

(ii) state the *size* of the resulting matrix products.

Exercise 6

Form all the possible matrix products from the matrices

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 3 & 2 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & -2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

Exercise 7

Given the matrices

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 1 \\ -1 & 2 & 1 \\ 3 & 4 & -2 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 2 & 3 \\ 3 & 4 & 1 \end{bmatrix},$$

find the matrix products **AB** and **BA**.

[*Solution to Exercises 5–7 on p. 56*]

The ideas of this subsection can be summarized in the following definition of matrix multiplication.

> The matrix $\mathbf{C} = \mathbf{AB}$ can be formed if the number of columns of the left-hand matrix **A** equals the number of rows of the right-hand matrix **B**. The matrix **C** will have the same number of rows as **A**, and the same number of columns as **B**, and its elements are given by
>
> $$\begin{aligned} c_{ij} &= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \\ &= \sum_{k=1}^{n} a_{ik}b_{kj} \end{aligned}$$
>
> where n is the number of columns of **A**.

2.3 The algebraic laws governing matrix multiplication

Before we can freely use matrix multiplication, we need to know whether matrix multiplication is commutative and associative, and whether there is a distributive law.

As you can see from the results of Exercise 7, **AB** and **BA** are not the same. Any counter-example would have been adequate to show that, in general, matrix multiplication is not commutative, that is,

> **AB** is not usually equal to **BA**.

In fact, we could have seen this without doing an example at all. If we multiply an $m \times n$ matrix **A** to an $n \times p$ matrix **B**, we obtain an $m \times p$ matrix **C**. However, it is not even possible to form the product **BA** unless p is equal to m.

As this is the first thing that we have discovered about matrices which differs significantly from real number algebra, we shall have to be very careful.

For instance, if

$$\mathbf{A} = \mathbf{B} \tag{1}$$

and we wish to multiply both sides of this matrix equation by another matrix **C**, then we have to decide either to multiply both sides of (1) by **C** on the left, to obtain

$$\mathbf{CA} = \mathbf{CB}, \tag{2}$$

or we could (size permitting) multiply both sides of (1) on the right, to obtain

$$\mathbf{AC} = \mathbf{BC}. \tag{3}$$

Both the matrix equations (2) and (3) would be logical consequences of (1), whereas there would be no justification in assuming that that **AC** and **CB** were equal.

Despite this bad beginning to the investigation of the laws governing matrix multiplication, in fact the associative and distributive laws are exactly the same as those in real number algebra. However, as matrix multiplication is not commutative, there are two versions of the distributive law to prove:

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$

and

$$(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}.$$

The proofs of these, and of the associative law

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$$

are given as answers to the following exercise. They are there for completeness, but please don't attempt to do them unless you are very brave, and have time on your hands! Exercises 16 and 17 at the end of the section will demonstrate these laws for three particular matrices. You are strongly advised to do these exercises now if you haven't had much practice at matrix multiplication before.

A consequence of the associative law, $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$, is that we can omit the brackets where appropriate and write **ABC** without fear of ambiguity.

Exercise 8 (*Optional*)

Assume (for simplicity) that **A**, **B** and **C** are all $n \times n$ square matrices. Prove, using the definition of matrix multiplication given in the box on p. 14, that:

(i) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$,

(ii) $(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$,

(iii) $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$.

[*Solution on p. 56*]

These laws, along with the laws of Section 1 and the inverse matrix in Section 4 (the matrix equivalent of the number a^{-1}), will enable us to do algebra with

matrices in much the same kind of way as we do with real numbers. The only difference will be that we shall have to be careful not to interchange expressions like **AB** and **BA**.

2.4 The product of a matrix and a column vector

The product of a matrix and a column vector produces another column vector—for example,

$$\begin{bmatrix} 2 & 7 & 1 \\ 3 & 4 & 2 \\ 5 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 15 \\ 9 \\ 6 \end{bmatrix}.$$

Exercise 9

If $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix}$ and $\mathbf{q} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, find $\mathbf{Aq}$.

[*Solution on p. 57*]

Example 5

If $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, then

$$\mathbf{Ax} = \begin{bmatrix} 3x_1 + x_2 \\ 4x_1 + 2x_2 \end{bmatrix}.$$

In this last example, the column vector formed looks very like the left-hand side of a pair of simultaneous equations. It can be seen that any matrix **A** multiplied by an 'unknown' column vector **x** provides us with a column vector whose elements are linear combinations of the elements of **x**. This enables us to write any set of linear simultaneous equations in the form

$$\mathbf{Ax} = \mathbf{b}.$$

Example 6

Given that $\quad \mathbf{A} = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 5 \end{bmatrix},$

write down the simultaneous equations represented by

$$\mathbf{Ax} = \mathbf{b}.$$

Solution

$$\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$$

gives us

$$\begin{bmatrix} x_1 - x_2 \\ 2x_1 + 3x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}.$$

By definition of equality of matrices, this gives us the simultaneous equations

$$\begin{aligned} x_1 - x_2 &= 0 \\ 2x_1 + 3x_2 &= 5. \end{aligned}$$

Exercise 10

Write down the simultaneous equations represented by $\mathbf{Ax} = \mathbf{b}$

if $\mathbf{A} = \begin{bmatrix} 3 & 1 & -2 \\ 1 & 2 & 1 \\ 4 & 1 & 0 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}.$

[*Solution on p. 57*]

Example 7

Express the equations

$$\begin{aligned} 2x_1 + \ x_2 &= 3 \\ x_1 - 3x_2 &= 5 \end{aligned}$$

in the form $\mathbf{Ax} = \mathbf{b}$.

Solution

Reversing the procedure of Example 6, the left-hand side of these equations can be written as

$$\begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Note that A is the matrix of the left-hand side coefficients.

So the equations above can be written as

$$\begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

which has the required form $\mathbf{Ax} = \mathbf{b}$.

Exercise 11

Write the equations

$$\begin{aligned} 2x_1 - 3x_2 + x_3 &= 1 \\ 7x_1 + x_2 - x_3 &= 2 \\ 9x_1 + 3x_2 + 5x_3 &= -1 \end{aligned}$$

in the form $\mathbf{Ax} = \mathbf{b}$.

[*Solution on p. 57*]

The reason that this notation was not used in the unit on simultaneous equations was that the operation of matrix multiplication had not been discussed at that time.

There is another context in which linear combinations and matrices arise. Suppose that we have two pairs of numbers (x_1, y_1) and (x_2, y_2) such that

$$\left.\begin{aligned} x_2 &= x_1 + 2y_1 \\ y_2 &= 4x_1 + y_1 \end{aligned}\right\} \qquad (1)$$

These equations allow us to put values of x_1 and y_1 in the right-hand side of (1), and hence calculate the values of x_2 and y_2. In a sense, we have 'transformed' (x_1, y_1) into (x_2, y_2), so a set of equations, like (1) above, is called a **transformation**—in this case a **linear transformation.** All linear transformations may be written in matrix form. For instance, the Equations (1) above can be written as

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}.$$

Thinking of transformations in this matrix form enables us to manipulate them more easily. For instance, suppose we know that

$$\left.\begin{aligned} x_2 &= x_1 + 2y_1 \\ y_2 &= 4x_1 + y_1 \end{aligned}\right\} \ (1) \qquad \text{and} \qquad \left.\begin{aligned} x_3 &= 3x_2 - y_2 \\ y_3 &= x_2 + 2y_2 \end{aligned}\right\} \ (2),$$

and we wish to express x_3 and y_3 in terms of x_1 and y_1.

We know that (1) and (2) above can be expressed in matrix form as

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \ (3) \qquad \text{and} \qquad \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \ (4).$$

Then we can replace the expression for the column vector $\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$ from (3) into (4),

to obtain

$$\begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 1 & 2 \end{bmatrix} \left\{ \begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \right\}$$
$$= \left\{ \begin{bmatrix} 3 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix} \right\} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

using the associative law for matrices. Thus x_3 and y_3 can be expressed more simply in terms of x_1 and y_1 as

$$\begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} -1 & 5 \\ 9 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}.$$

Rather more generally, suppose we perform a sequence of linear transformations of this kind:

$$\mathbf{r}_2 = \mathbf{A}\mathbf{r}_1 \quad (3) \qquad \mathbf{r}_3 = \mathbf{B}\mathbf{r}_2 \quad (4) \qquad \text{where } \mathbf{r}_k = \begin{bmatrix} x_k \\ y_k \end{bmatrix}.$$

Then by substituting the expression for $\mathbf{r}_2$ in (3) into (4), we get

$$\mathbf{r}_3 = \mathbf{B}(\mathbf{A}\mathbf{r}_1) = (\mathbf{B}\mathbf{A})\mathbf{r}_1,$$

and hence we obtain expressions for x_3 and y_3 in terms of x_1 and y_1.
You will need this concept of combining successive matrix transformations using the multiplication process to understand what is happening in the television programme.

Exercise 12

If $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 3 & 1 \\ 0 & -4 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$ and $\mathbf{r}_k = \begin{bmatrix} x_k \\ y_k \end{bmatrix}$,

express the following succession of matrix transformations:

$$\mathbf{r}_2 = \mathbf{A}\mathbf{r}_1, \quad \mathbf{r}_3 = \mathbf{B}\mathbf{r}_2 \quad \text{and} \quad \mathbf{r}_4 = \mathbf{C}\mathbf{r}_3$$

as a single matrix transformation

$$\mathbf{r}_4 = \mathbf{Q}\mathbf{r}_1.$$

Find the matrix $\mathbf{Q}$, first in terms of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and then as a matrix with numerical elements.
[*Solution on p. 57*]

2.5 The matrix transpose

> The **transpose** of a matrix $\mathbf{A}$ (written as $\mathbf{A}^T$ in this course, or sometimes as $\mathbf{A}'$) is simply the matrix whose columns are the rows of $\mathbf{A}$.

For instance, if

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 3 \\ 0 & 4 & 5 \end{bmatrix},$$

then

$$\mathbf{A}^T = \begin{bmatrix} 2 & 0 \\ 1 & 4 \\ 3 & 5 \end{bmatrix}.$$

You can see that the first row of $\mathbf{A}$ becomes the first column of $\mathbf{A}^T$ and the second row of $\mathbf{A}$ becomes the second column of $\mathbf{A}^T$.

Exercise 13

If $\mathbf{A} = \begin{bmatrix} 2 & 3 & 7 \\ 1 & 0 & 2 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \\ -1 & 7 \end{bmatrix}$,

write down $\mathbf{A}^T$ and $\mathbf{B}^T$.

Exercise 14

Show for the matrices in Exercise 13 that

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T.$$

[*Solutions to Exercises 13 and 14 on p. 57*]

The result

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T \tag{5}$$

is always true, although it is not proved in this unit. We shall assume this result and show that a similar result is true for three matrices.

Example 8

To show that $(\mathbf{ABC})^T = \mathbf{C}^T\mathbf{B}^T\mathbf{A}^T$.

Proof

$$\begin{aligned}(\mathbf{ABC})^T &= (\mathbf{A}(\mathbf{BC}))^T \\ &= (\mathbf{BC})^T\mathbf{A}^T && \text{using (5)} \\ &= \mathbf{C}^T\mathbf{B}^T\mathbf{A}^T && \text{using (5)}\end{aligned}$$

This is just another example of how to use matrix algebra.

The general result is

$$\boxed{(\mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_n)^T = \mathbf{A}_n^T\mathbf{A}_{n-1}^T\ldots\mathbf{A}_2^T\mathbf{A}_1^T.}$$

An example of the use of a matrix transpose is in the theoretical matrix description of the linear programming problem in *Unit 10.* A typical linear programming problem would be:

Minimize $c_1x_1 + c_2x_2 + c_3x_3$

subject to the conditions

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 \qquad\qquad &\le b_1 \\ a_{21}x_1 \qquad\qquad + a_{23}x_3 &\le b_2 \\ x_1, x_2, x_3 \ge 0.\end{aligned}$$

In matrix terms, this can be expressed as:

Minimize $\begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

subject to the conditions

$$\begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & 0 & a_{23} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \le \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

$$x_1, x_2, x_3 \ge 0.$$

Rather more briefly, we say:

Minimize $\mathbf{c}^T\mathbf{x}$

subject to the conditions

$\mathbf{Ax} \le \mathbf{b}$

and $\mathbf{x} \ge \mathbf{0}$.

We are forced to use the notation $\mathbf{c}^T$ as, to save confusion, we stated at the beginning of the unit that the notation $\mathbf{c}$ refers only to a column vector. So if we need a row vector, as in this case, we use the notation $\mathbf{c}^T$.

Finally, note the economy in expressing our linear programming problem in terms of matrices. You will find that matrix notation gives us an excellent short way to express and to simplify problems.

Summary of Section 2

1. The matrix $\mathbf{C} = \mathbf{AB}$ can be formed if the number of columns of the left-hand matrix $\mathbf{A}$ equals the number of rows of the right-hand matrix $\mathbf{B}$.

The product matrix $\mathbf{C}$ will have the same number of rows as $\mathbf{A}$, and the same number of columns as $\mathbf{B}$, and its elements are given by

$$\begin{aligned} c_{ij} &= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \\ &= \sum_{k=1}^{n} a_{ik}b_{kj} \end{aligned}$$

where n is the number of columns of $\mathbf{A}$.

2. The commutative law for matrix multiplication is *not* true. So, in general, $\mathbf{AB}$ is not usually equal to $\mathbf{BA}$. However, the multiplicative associative law

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$$

and the distributive laws

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$

and

$$(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$$

are true.

3. The product of a matrix and a column vector is another column vector. This enables us to express:

(i) simultaneous equations in the form

$$\mathbf{Ax} = \mathbf{b};$$

(ii) **linear transformations** in the form

$$\mathbf{r}_2 = \mathbf{A}\mathbf{r}_1, \text{ where } \mathbf{r}_k = \begin{bmatrix} x_k \\ y_k \end{bmatrix}.$$

4. The **transpose** of a matrix $\mathbf{A}$ is the matrix $\mathbf{A}^T$ whose columns are the rows of $\mathbf{A}$. The main property of transposed matrices is

$$(\mathbf{A}_1\mathbf{A}_2 \ldots \mathbf{A}_n)^T = \mathbf{A}_n^T\mathbf{A}_{n-1}^T \ldots \mathbf{A}_2^T\mathbf{A}_1^T.$$

End of section exercises

There are more end of section exercises than usual. This is to ensure that those of you who haven't met matrix multiplication before get adequate practice.

Exercise 15

Given the matrices

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 1 \\ 2 & 3 & 1 \end{bmatrix},$$

form the following products where possible:

(i) $\mathbf{Ab}$ (ii) $\mathbf{AC}$ (iii) $\mathbf{AD}$ (iv) $\mathbf{bC}$ (v) $\mathbf{b}^T\mathbf{C}$ (vi) $\mathbf{DC}^T$.

Exercise 16

Show that the associative law $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ and the two versions of the distributive law, $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ and $(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$, hold for the matrices

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 1 \\ -1 & 2 & 1 \\ 3 & 4 & -2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 2 & 3 \\ 3 & 4 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix}.$$

(You have already found $\mathbf{AB}$ and $\mathbf{BA}$ in Exercise 7.)

Exercise 17

Show that the distributive laws stated in the last exercise are true for the matrices

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 7 \\ 1 & 0 & -2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \\ 5 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 4 & -1 \end{bmatrix}.$$

Why is the associative law not applicable?

Exercise 18

Express the simultaneous equations

$$\begin{aligned} 2x_1 + 3x_2 + 4x_3 &= -1 \\ x_1 - 2x_2 + 5x_3 &= 0 \\ x_2 - 2x_3 &= 10 \end{aligned}$$

in the form $\mathbf{Ax} = \mathbf{b}$.

Exercise 19

Combine the transformations

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

into a single transformation, giving $\begin{bmatrix} x_3 \\ y_3 \end{bmatrix}$ in terms of $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$.

[*Solutions to Exercises 15–19 on pp. 57–58*]

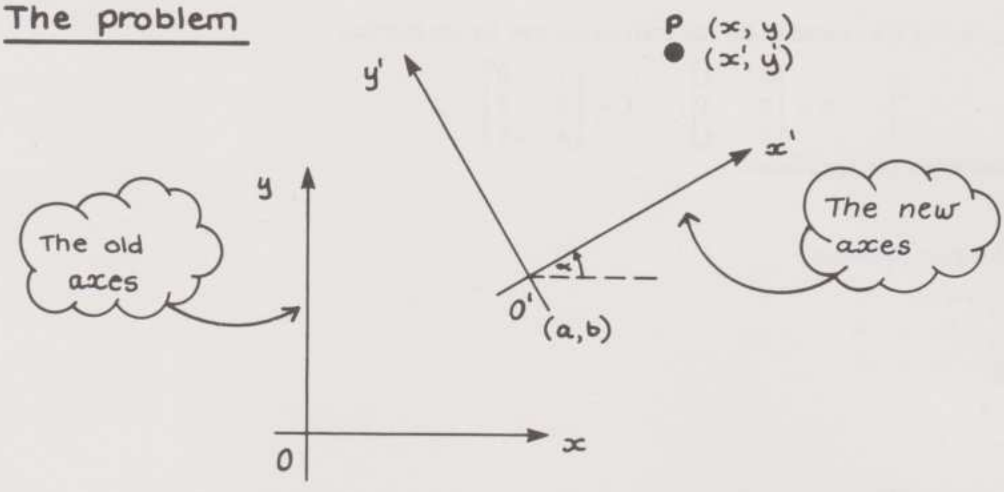
3 Change of axes

3.1 Change of axes in two dimensions (Tape Subsection)

This subsection is a tape subsection in preparation for the television programme.
Start the tape now.

1

The problem



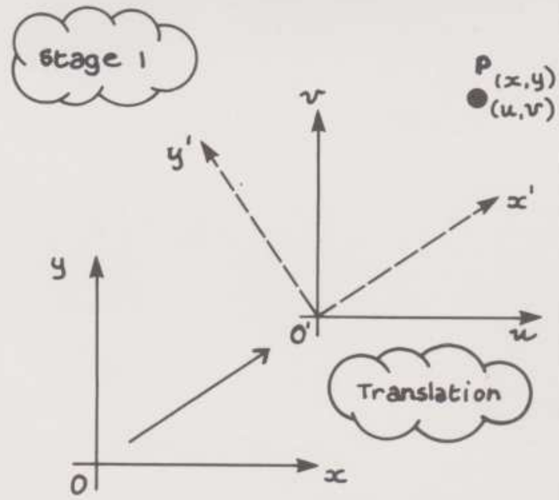
Given the co-ordinates (x, y) of a point P with respect to some axes, what are the co-ordinates (x', y') of this point with respect to some new axes?

2

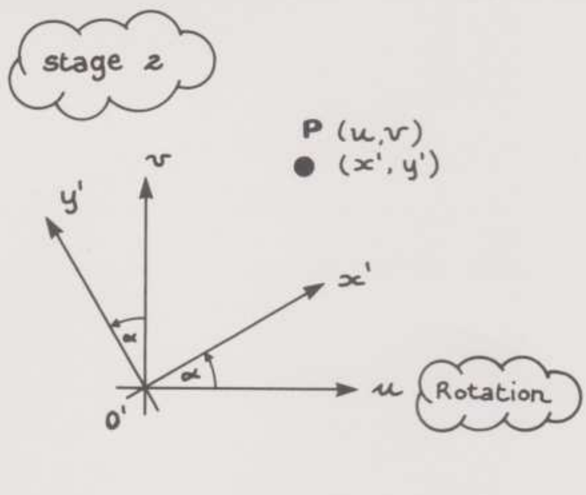
The strategy

Stage 0 — Split the problem up

Stage 1



Stage 2

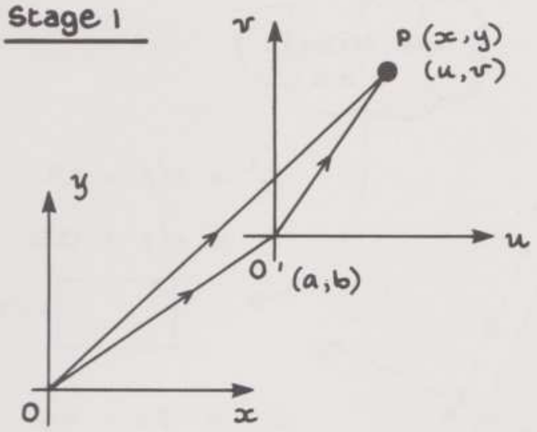


Stage 3 — Put stages 1 and 2 together:

$$(x, y) \xrightarrow[\text{Stage 1}]{\text{Translation}} (u, v) \xrightarrow[\text{Stage 2}]{\text{Rotation}} (x', y')$$

3

Stage 1

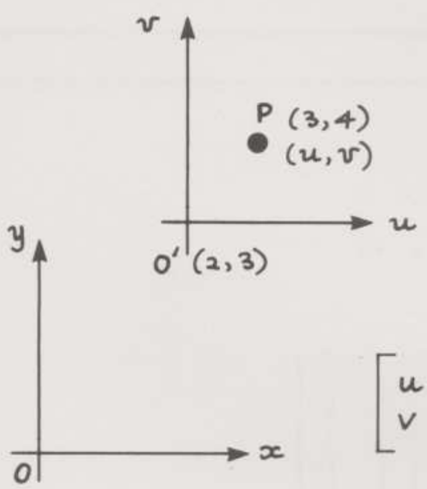


$$\overrightarrow{O'P} = \overrightarrow{OP} - \overrightarrow{OO'}$$

$$\boxed{\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \quad \\ \quad \end{bmatrix} - \begin{bmatrix} \quad \\ \quad \end{bmatrix}}$$

4

Exercise 1



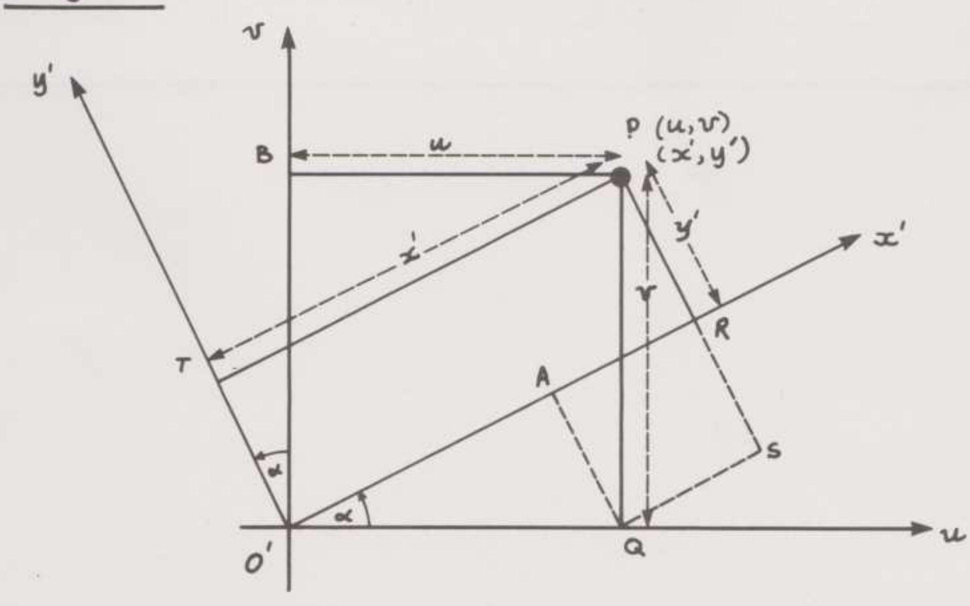
What are the co-ordinates of P with respect to the O'uv axes?

Answer in the back of the text

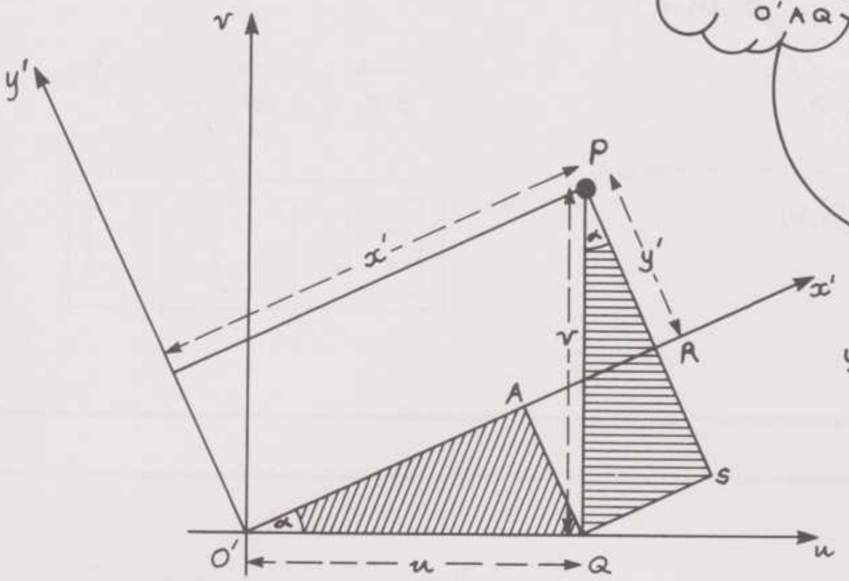
$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} a \\ b \end{bmatrix}, \text{ so } \quad \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \quad \\ \quad \end{bmatrix}$$

5

Stage 2



6 Rotation



use triangle O'AQ

$$x' = O'A + AR$$

$$= O'A + QS$$

$$= \boxed{} + v \sin\alpha$$

$$y' = PS - RS$$

$$= \boxed{} - AQ$$

$$= \boxed{}$$

7 The rotation matrix

$$x' = u \cos\alpha + v \sin\alpha$$

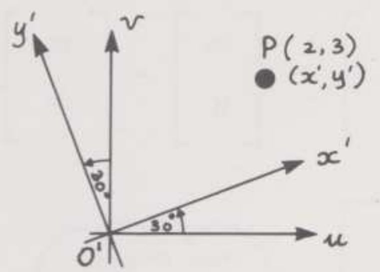
$$y' = -u \sin\alpha + v \cos\alpha$$

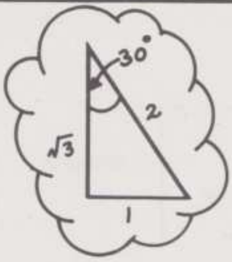
The equivalent matrix equation:

$$\boxed{\begin{bmatrix} x' \\ y' \end{bmatrix} \begin{matrix} = \\ = \end{matrix} \begin{bmatrix} & \\ & \end{bmatrix} \begin{bmatrix} \\ \end{bmatrix}}$$

[The new Point] is produced by [the rotation matrix] operating on [the old point]

8 Exercise 2





What are the coordinates of P with respect to the $O'x'y'$ axes?

Answer in the back of text

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

$$= \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix} \begin{bmatrix} \quad \\ \quad \end{bmatrix}$$

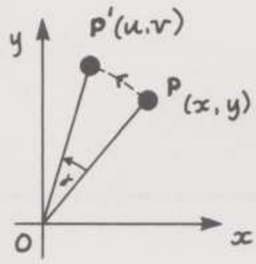
$$= \begin{bmatrix} \quad \\ \quad \end{bmatrix}$$

so $x' = \boxed{\quad}$

$y' = \boxed{\quad}$

9 Different point of view

(a)

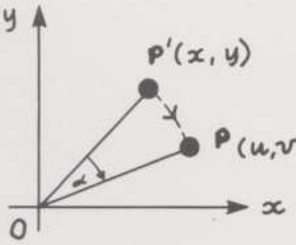


m101 style

(a) P' has coordinates (u, v), where

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

(b)



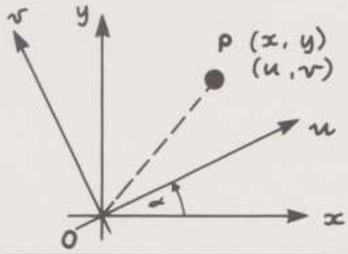
m101 style

(b) P' has coordinates (u, v) where

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

(c)



style in this unit

(c) The new coordinates (u, v) of P are

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

10 **Stage 3**

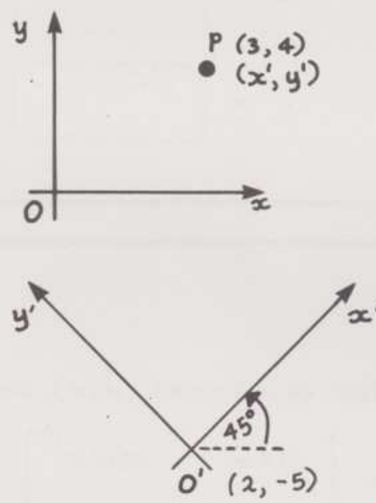
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \quad \text{where} \quad \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} a \\ b \end{bmatrix}$$

Thus

$$\boxed{\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix} \begin{bmatrix} \quad \\ \quad \end{bmatrix}}$$

The answer in terms of x and y

11 **Exercise 3**



What are the coordinates of P relative to the $O'x'y'$ axes?

Answer in the back of text

3.2 The re-orientation problem (Television Subsection)

TV20

Read the following notes before viewing the television programme.

In the television programme, we solve a three-dimensional problem which is the equivalent of the two-dimensional change of axes problem in the tape subsection. In Figure 1 below, for instance, we require to find the new coordinates (x', y', z') of P with respect to an origin O', given that P has coordinates (x, y, z) with respect to the origin O.

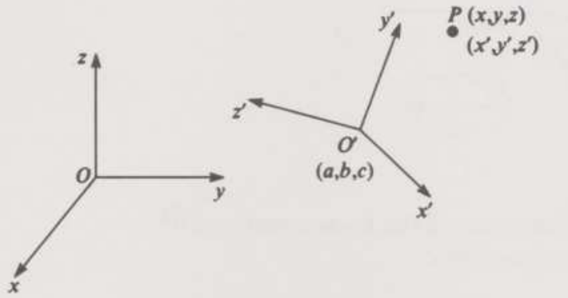


Figure 1

In the television programme, O is the planet Earth, P is the planet Saturn, and O' is the space probe Voyager. The displacement vector $\overrightarrow{OP}$ of Saturn relative to Earth is referred to as $\mathbf{s}$, and the displacement vector $\overrightarrow{OO'}$ of Voyager relative to Earth is referred to as $\mathbf{v}$. Thus from *Unit 14* on vector algebra, we know that the displacement vector $\overrightarrow{O'P}$ of Saturn relative to Voyager will be $\mathbf{s} - \mathbf{v}$. These vectors are marked in Figure 2.

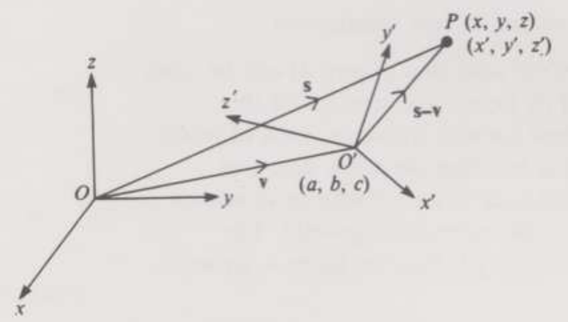


Figure 2

At the beginning of the programme, you will see an explicit set of equations which determine the new coordinates, written without the use of matrices. You will see that these are very long and messy to write down, and have no obvious structure. The programme derives a matrix solution to the problem which is (a) more readily understandable, (b) easy to derive from first principles, and (c) allows us to see some structure in the resulting solution.

When you watch the programme, you should concentrate on the strategy, rather than on the detail of the matrices. These are all given in the post-programme notes.

Now watch the television programme: 'Applying matrices—the algebra of re-orientation'.

Read the following notes after watching the programme. (They summarize the programme, and could help to serve as a substitute for it if you have not been able to view it.)

The problem in the programme is the three-dimensional equivalent of the problem on the tape. (See Figure 1 above.)

P (the planet Saturn) has a known set of coordinates with respect to a set of axes $Oxyz$, centred on Earth. Given another set of axes of known orientation whose origin is at a known point O' (the space probe Voyager), what are the coordinates of P with respect to these other axes?

As before, the problem can be broken down into stages.

Stage 1

It is a relatively easy matter to get the coordinates of P with respect to axes at O' which are parallel to the original axes. We call these intermediate coordinates u, v, w.

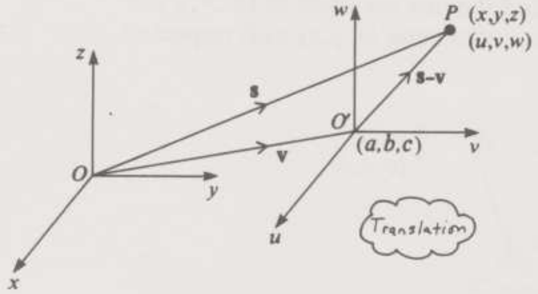


Figure 3

These can be found by finding the vector $\overrightarrow{O'P}$ in terms of the known vectors $\overrightarrow{OP}$ and $\overrightarrow{OO'}$. Using the vector rule of addition, we know that

$$\overrightarrow{O'P} = \overrightarrow{OP} - \overrightarrow{OO'}.$$

In terms of column vectors, this gives us

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} a \\ b \\ c \end{bmatrix}. \tag{1}$$

Stage 2

The rotation is not the simple problem that we had in two dimensions.

The alignment of the $O'uvw$ axes with the $O'x'y'z'$ axes (see Figure 4) can be done in various ways. The way we use is attributed to Euler. The strategy of this method is to break the rotation down into three simpler rotations, each of which holds one of the coordinate axes fixed. The first rotation we make does not actually align any of the axes, but positions the axes in such a way as to ensure that the two subsequent rotations will perform the entire re-alignment. The advantage of breaking the problem down in this way is that we know a formula for two-dimensional rotations.

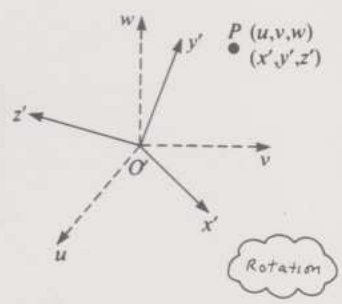


Figure 4

Imagine that the w-axis in Figure 5 is sticking up vertically into the air, and the uv-plane is the ground. If a rotation is made, keeping the w-axis fixed, the rotation will not change the w coordinate, as the distance of the point above the ground will not change. This, and the results from the tape subsection, give us

$$\begin{aligned} u' &= u\cos\alpha + v\sin\alpha \\ v' &= -u\sin\alpha + v\cos\alpha \\ w' &= w. \end{aligned}$$

In matrix terms, we get the coordinates of P with respect to the axes $O'u'v'w'$ as

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}. \tag{2}$$

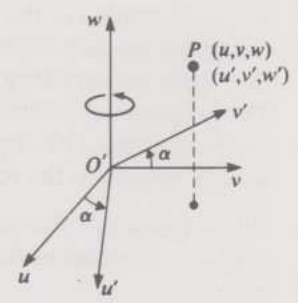


Figure 5

If we do a subsequent rotation, this time keeping the u'-axis fixed, and rotating through an angle β in the $v'w'$-plane, then Figure 6 shows that the u' coordinate will not change, and we get yet another set of coordinates for P, with respect to another set of axes, $O'u''v''w''$:

$$\begin{aligned} u'' &= u' \\ v'' &= v'\cos\beta + w'\sin\beta \\ w'' &= -v'\sin\beta + w'\cos\beta. \end{aligned}$$

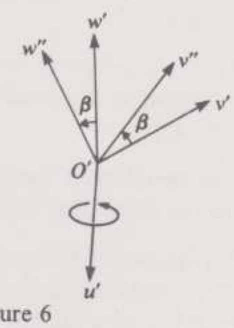


Figure 6

This can be expressed in matrix terms as

$$\begin{bmatrix} u'' \\ v'' \\ w'' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & \sin\beta \\ 0 & -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix}. \tag{3}$$

If yet a third rotation is done, keeping w'' fixed, and rotating about the w''-axis through an angle γ, then the w'' coordinate will not change. Similarly to the first rotation, we obtain the coordinates of P with respect to yet another set of axes $O'u'''v'''w'''$:

$$\begin{bmatrix} u''' \\ v''' \\ w''' \end{bmatrix} = \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u'' \\ v'' \\ w'' \end{bmatrix}. \tag{4}$$

Suitable choice of the angles α, β and γ (sometimes called the **Eulerian angles**) will ensure that the $O'u'''v'''w'''$ axes thus obtained are the required $O'x'y'z'$ axes, and thus that the coordinates (u''', v''', w''') are the required coordinates (x', y', z'). How these angles are chosen is shown in the programme, and illustrated in Figure 7 below. (To avoid a confusing figure, the v' and v'' axes are not shown.)

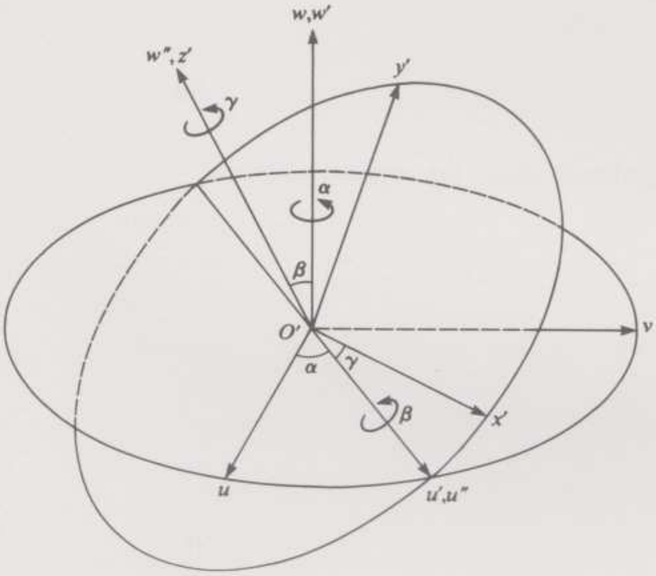


Figure 7

The uv-plane and the $x'y'$-plane will intersect in a straight line. (Figure 7 shows two circular portions of these planes, centred at the origin, intersecting in the line marked $O'u'$.) The three rotations are carried out as follows.

Step 1

Rotate about the w-axis until the u-axis is in the line of intersection of the uv- and $x'y'$-planes. The α is the angle of this rotation. (The new u-axis is called the u'-axis.)

Step 2

Rotate about the u'-axis until the w'-axis coincides with the z'-axis. Then β is the angle of this rotation.

Step 3

Rotate about the w''-axis until the u''- and v''-axes coincide with the x'- and y'-axes respectively. Then γ is the angle of this rotation.

We are now in a position to find an expression for the coordinates (x', y', z') in terms of the coordinates (u, v, w). We know that

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \mathbf{M}_\gamma \begin{bmatrix} u'' \\ v'' \\ w'' \end{bmatrix} \tag{4}$$

where

$$\begin{bmatrix} u'' \\ v'' \\ w'' \end{bmatrix} = \mathbf{M}_\beta \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} \tag{3}$$

and

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \mathbf{M}_\alpha \begin{bmatrix} u \\ v \\ w \end{bmatrix} \tag{2}$$

and $\mathbf{M}_\alpha$, $\mathbf{M}_\beta$ and $\mathbf{M}_\gamma$ are the rotation matrices on pp. 28–29.

Combining these transformations in the way we did in the last section, and using the associative rule, we obtain

$$\begin{aligned} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} &= \mathbf{M}_\gamma \mathbf{M}_\beta \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} && \text{from (3)} \\ &= \mathbf{M}_\gamma \mathbf{M}_\beta \mathbf{M}_\alpha \begin{bmatrix} u \\ v \\ w \end{bmatrix} && \text{from (2)} \\ &= (\mathbf{M}_\gamma \mathbf{M}_\beta \mathbf{M}_\alpha) \begin{bmatrix} u \\ v \\ w \end{bmatrix}. && (5) \end{aligned}$$

Thus we can multiply $\mathbf{M}_\gamma$, $\mathbf{M}_\beta$ and $\mathbf{M}_\alpha$ together to obtain the required relation between x', y', z' and u, v, w in the form

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \mathbf{M} \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

where

$$\mathbf{M} = \mathbf{M}_\gamma \mathbf{M}_\beta \mathbf{M}_\alpha.$$

Finally, we know that

$$\begin{aligned} \begin{bmatrix} u \\ v \\ w \end{bmatrix} &= \begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} a \\ b \\ c \end{bmatrix} && (1) \\ &= \begin{bmatrix} x - a \\ y - b \\ z - c \end{bmatrix}, \end{aligned}$$

where (x, y, z) are the coordinates of P relative to the original origin O. So, substituting this into the matrix equation (5) above, we obtain the new coordinates for the point P in terms of the old coordinates:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \mathbf{M}_\gamma \mathbf{M}_\beta \mathbf{M}_\alpha \begin{bmatrix} x - a \\ y - b \\ z - c \end{bmatrix}. \tag{6}$$

Post-television work

In the following exercise we ask you to do the final step yourself, thus returning full circle to the beginning of the television programme.

Exercise 4

Obtain three explicit equations, giving x', y' and z' in terms of x, y, z, the displacement coordinates a, b, c, and the Eulerian angles α, β, γ.

[*Solution on p. 59*]

End of television programme notes

Summary of Section 3

1. Given that the coordinates of P are (x, y) with respect to some axes Ox, Oy with their origin at O, then the coordinates of P with respect to another pair of axes $O'x'$, $O'y'$ with their origin at (a, b) and inclined at an anticlockwise angle α to the original axes (see Figure 8) are given by

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x-a \\ y-b \end{bmatrix}.$$

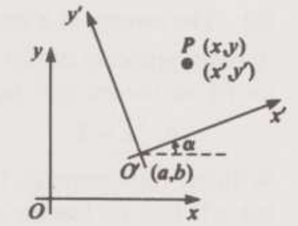


Figure 8

2. In three dimensions, if we wish to move the origin to the point (a, b, c) and re-orientate the axes (see Figure 9) the new coordinates of P are given by

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \mathbf{M}_\gamma \mathbf{M}_\beta \mathbf{M}_\alpha \begin{bmatrix} x-a \\ y-b \\ z-c \end{bmatrix}$$

where α, β and γ are the **Eulerian angles** of the rotation and $\mathbf{M}_\alpha$, $\mathbf{M}_\beta$, $\mathbf{M}_\gamma$ are 3×3 rotation matrices each describing a rotation which holds one of the intermediate coordinate axes fixed.

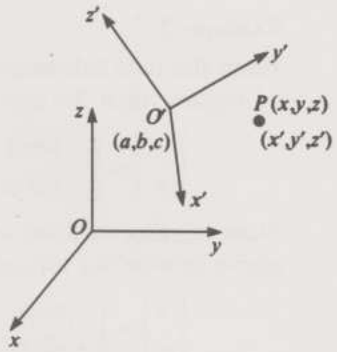


Figure 9

4 Square matrices and their inverses

4.1 Square matrices

In this section, we will restrict ourselves to the algebra of square matrices. The reason for doing this is that:

(i) there are important results which are true only for square matrices, and

(ii) real-life situations often produce square matrix blocks.

Square matrices of the same size can be both added and multiplied, and thus satisfy all the laws of algebra discussed in the first part of this unit. Also, suitably sized zeros ($\mathbf{0}$) and negative elements ($-\mathbf{A}$) are already defined. But before we can use matrix algebra in a way similar to real number algebra, we still need to find a matrix analogous to the number 1, and another matrix analogous to the number a^{-1}. This section discusses these matrices, and the resulting matrix algebra.

(a) The unit (or identity) matrix

This is a square matrix whose diagonal elements $a_{11}, a_{22} \ldots a_{nn}$ are all ones, and all other elements are zeros. (The ones are said to lie on the **main diagonal**, which starts in the top left-hand corner.) Examples of unit matrices are

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \text{ and } \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Matrices such as these are referred to as $\mathbf{I}$. The size of $\mathbf{I}$ is not normally specified, as this is usually clear from the context.

Unit matrices have the important property that, for any square matrix $\mathbf{A}$,

$$\boxed{\mathbf{AI} = \mathbf{IA} = \mathbf{A}.}$$

The following example demonstrates this property for one particular matrix.

Example 1

If $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 7 & 8 \end{bmatrix}$, then

$$\mathbf{AI} = \begin{bmatrix} 2 & 1 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2+0 & 0+1 \\ 7+0 & 0+8 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 7 & 8 \end{bmatrix} = \mathbf{A}$$

and

$$\mathbf{IA} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 2+0 & 1+0 \\ 0+7 & 0+8 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 7 & 8 \end{bmatrix} = \mathbf{A}.$$

(b) The inverse of a square matrix A

A basic property of our real number system is that for each real number a (except for 0), an inverse a^{-1} exists so that

$$a^{-1}a = 1.$$

Without this property, it would be impossible to solve equations. This leads us to ask whether an **inverse matrix** $\mathbf{A}^{-1}$ exists so that

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

Example 2

From the tape subsection, we know that if the x- and y-axes are rotated through an angle α, then the new coordinates of a point P will be given by

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \qquad (1)$$

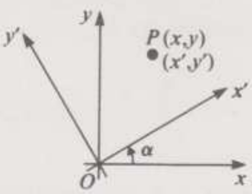


Figure 1

Now, starting with the co-ordinates (x', y'), we know we can 'undo' the process above by rotating through a clockwise angle α. So we know that

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos(-\alpha) & \sin(-\alpha) \\ -\sin(-\alpha) & \cos(-\alpha) \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$
$$= \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

Substituting this expression for $\begin{bmatrix} x \\ y \end{bmatrix}$ into (1), we get

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

Thus the matrix product

$$\begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

leaves all vectors $\begin{bmatrix} x' \\ y' \end{bmatrix}$ fixed and must therefore be the identity matrix, **I**. So, if we call the original rotation matrix **M**, then the matrix which 'undoes' the rotation must be $\mathbf{M}^{-1}$.

Exercise 1

Given that

$$\mathbf{M} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \quad \text{and} \quad \mathbf{N} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix},$$

verify by multiplying the matrices together that $\mathbf{MN} = \mathbf{NM} = \mathbf{I}$.

Exercise 2

Given that

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

where $ad - bc \neq 0$, then verify that

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}.$$

[*Solutions to Exercises 1 and 2 on p. 59*]

From this last exercise, it can be seen that if $ad - bc$ is not zero, the 2×2 matrix

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

has an inverse

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \qquad (2)$$

such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

> When the inverse of an $n \times n$ matrix exists,
>
> $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ (3)

In 2×2 matrices where $ad - bc$ is zero, then no inverse exists; but whenever a matrix $\mathbf{A}$ has an inverse $\mathbf{A}^{-1}$ then the commutative property (3) is true. This will be proved more generally in Subsection 4.4.

Example 3

Using Equation (2), the inverse of $\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$ is $\frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$, but $\begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ have no inverses, as, in both cases, $ad - bc = 0$.

Exercise 3

Using Equation (2), write down the inverses (if possible) of the following matrices:

(i) $\begin{bmatrix} 2 & 7 \\ 1 & 6 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}$ (iii) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

[*Solution on p. 59*]

One use of the inverse matrix is in the theoretical simplification of matrix expressions, and in the theoretical solution of matrix equations.

If we are given a real number equation

$$ax = b,$$

then the theoretical solution of this equation can be written

$$x = a^{-1}b.$$

Similarly, if we are given a set of simultaneous equations, represented by the matrix equation

$$\mathbf{Ax} = \mathbf{b}, \tag{4}$$

then we can also find a theoretical solution of this equation in the following way.

Assuming that $\mathbf{A}^{-1}$ exists, we can multiply both sides of Equation (4) on the left by $\mathbf{A}^{-1}$ to give

$$\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{b}.$$

So

$$\mathbf{Ix} = \mathbf{A}^{-1}\mathbf{b} \qquad \text{as } \mathbf{A}^{-1}\mathbf{A} = \mathbf{I},$$

i.e.

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \qquad \text{as } \mathbf{Ix} = \mathbf{x}.$$

It is in this kind of way that we use $\mathbf{A}^{-1}$ to obtain theoretical simplifications of matrix expressions, or theoretical solutions of matrix equations. In many problems, this is as far as we need go. However, the next subsection discusses a method for finding the numerical value of $\mathbf{A}^{-1}$, when this is necessary.

4.2 A method for finding $\mathbf{A}^{-1}$

From Subsection 4.1, we know that the theoretical solution of the simultaneous equations

$$\mathbf{Ax} = \mathbf{b} \tag{5}$$

is

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

So, a method for finding $\mathbf{A}^{-1}$ would be to solve the equations (5), as the following example demonstrates.

Example 4

Find the inverse of the matrix $\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$ by solving the equations

$$x_1 + 0x_2 = b_1 \qquad E_1$$
$$2x_1 + 3x_2 = b_2 \qquad E_2$$

Solution

Using Gaussian elimination we get

$$x_1 + 0x_2 = b_1 \qquad E_1$$
$$E_2 - 2E_1 \qquad 0x_1 + 3x_2 = b_2 - 2b_1 \qquad E_{2a}$$

This gives the solution

$$E_{2a} \div 3 \qquad x_2 = \tfrac{1}{3}b_2 - \tfrac{2}{3}b_1$$
$$x_1 = b_1.$$

In matrix terms, this means that the set of equations

$$\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},$$

which is of the form $\mathbf{Ax} = \mathbf{b}$, has the solution

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix},$$

which is of the form $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

So, by solving the equations $\mathbf{Ax} = \mathbf{b}$ we find that the inverse of $\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$.

As you know from *Unit 9*, the elimination process can be done by operations on the rows of the matrix $\mathbf{A}|\mathbf{b}$. In fact, the whole operation of solving the equations, back substitution included, can be done by matrix row operations very similar to those used in Gaussian elimination, as the following example demonstrates.

Example 5

Once again, we shall find the inverse of a 2×2 matrix, namely

$$\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix}.$$

This time, each step will be written down in matrix form.

Working	Matrix equivalent
To solve $x_1 + 2x_2 = b_1 \qquad E_1$ $x_1 + 4x_2 = b_2 \qquad E_2$	*To solve* $\quad \mathbf{A} \quad \mathbf{x} \quad = \quad \mathbf{I} \quad \mathbf{b} \qquad (5)$ $\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \begin{matrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{matrix}$ Note that **b** can be written as **Ib**.
The elimination $x_1 + 2x_2 = b_1 \qquad E_1$ $E_2 - E_1 \qquad 0x_1 + 2x_2 = b_2 - b_1 \qquad E_{2a}$	$\mathbf{R}_2 - \mathbf{R}_1 \quad \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \begin{matrix} \mathbf{R}_1 \\ \mathbf{R}_{2a} \end{matrix} \qquad (6)$

The matrix equation (6) is just a matrix form of equations E_1 and E_{2a}. However, with hindsight, we know we could have obtained both these matrices simply by subtracting $\mathbf{R}_1$ from $\mathbf{R}_2$, which is the exact equivalent of subtracting E_1 from E_2 in the equations.

To continue with the process of solving the equations:

Working	Matrix equivalent
First, find x_2. $x_1 + 2x_2 = b_1 \quad E_1$ $E_{2a} \div 2 \quad 0x_1 + x_2 = \frac{1}{2}(b_2 - b_1) \quad E_{2b}$	$\begin{array}{l} \\ \mathbf{R}_{2a} \div 2 \end{array} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \begin{array}{l} \mathbf{R}_1 \\ \mathbf{R}_{2b} \end{array}$

Again, the process of finding x_2 by dividing E_{2a} by two could have been achieved by dividing $\mathbf{R}_{2a}$ by two in both matrices.

Now, because we are limited to row operations, we make a slight departure from the normal procedure of back substitution. Instead of substituting the value we've found for x_2 into E_1 to find x_1, we subtract a suitable multiple of E_{2b} from E_1 to eliminate x_2 from the first equation. This is in fact exactly the same as back substitution, but it is arranged in a different way.

Working	Matrix equivalent
Now eliminate x_2 *from* E_1. $E_1 - 2E_{2b} \quad x_1 = b_1 - (b_2 - b_1) \quad E_{1a}$ $x_2 = \frac{1}{2}(b_2 - b_1) \quad E_{2b}$	$\begin{array}{l} \mathbf{R}_1 - 2\mathbf{R}_{2b} \\ \\ \end{array} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \begin{array}{l} \mathbf{R}_{1a} \\ \mathbf{R}_{2b} \end{array}$

Once again, it can be seen that the process of subtracting twice E_{2b} from E_1 could have been done directly by subtracting twice $\mathbf{R}_{2b}$ from $\mathbf{R}_1$ in the two matrices. In this example the process is now finished. (In a different example, we might still have had to divide E_{1a} (or $\mathbf{R}_{1a}$) by a factor to obtain x_1.) This gives us that the inverse of $\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix}$ is $\begin{bmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

The point of this example is that the entire operation of finding the inverse of a matrix can be done by row operations simultaneously performed on **A** and **I**.

> Starting with the matrices
>
> **A** and **I**,
>
> perform the same linear row operations simultaneously on these matrices. If **A** is 'reduced' to **I**, the same row operations will change **I** to $\mathbf{A}^{-1}$.

Thus, the example above could be done entirely in terms of matrix row operations as follows.

$$\begin{array}{l} \\ \\ \\ \\ \mathbf{R}_2 - \mathbf{R}_1 \\ \\ \mathbf{R}_{2a} \div 2 \\ \mathbf{R}_1 - 2\mathbf{R}_{2b} \\ \\ \\ \end{array}
\begin{array}{cc|cc}
\multicolumn{2}{c|}{\mathbf{A}} & \multicolumn{2}{c}{\mathbf{I}} \\
1 & 2 & 1 & 0 \\
1 & 4 & 0 & 1 \\
1 & 2 & 1 & 0 \\
0 & 2 & -1 & 1 \\
1 & 2 & 1 & 0 \\
0 & 1 & -\frac{1}{2} & \frac{1}{2} \\
1 & 0 & 2 & -1 \\
0 & 1 & -\frac{1}{2} & \frac{1}{2} \\
\multicolumn{2}{c|}{\mathbf{I}} & \multicolumn{2}{c}{\mathbf{A}^{-1}}
\end{array}
\begin{array}{l} \\ \mathbf{R}_1 \\ \mathbf{R}_2 \\ \mathbf{R}_1 \\ \mathbf{R}_{2a} \\ \mathbf{R}_1 \\ \mathbf{R}_{2b} \\ \mathbf{R}_{1a} \\ \mathbf{R}_{2b} \\ \\ \end{array}$$

It is advisable if doing these problems by hand to check that the matrix you have found really is the inverse, by calculating $\mathbf{A}^{-1}\mathbf{A}$ or $\mathbf{A}\mathbf{A}^{-1}$.

This method for finding the inverse is quite systematic, and like Gaussian elimination, it is suitable for implementation on a computer.

Example 6

Find $\mathbf{A}^{-1}$, given that

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 8 \\ 1 & 1 & 7 \end{bmatrix}.$$

Solution

Start with $\qquad \mathbf{A} \quad | \quad \mathbf{I}$

First get rid of the elements under the main diagonal using Gaussian elimination.

$$\begin{array}{r} \\ \\ \\ \end{array}\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 8 & 0 & 1 & 0 \\ 1 & 1 & 7 & 0 & 0 & 1 \end{array}\right]\begin{array}{l} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \mathbf{R}_3 \end{array}$$

$$\begin{array}{r} \\ \mathbf{R}_2 - 2\mathbf{R}_1 \\ \mathbf{R}_3 - \mathbf{R}_1 \end{array}\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & -1 & 4 & -1 & 0 & 1 \end{array}\right]\begin{array}{l} \mathbf{R}_1 \\ \mathbf{R}_{2a} \\ \mathbf{R}_{3a} \end{array}$$

$$\begin{array}{r} \\ \\ \mathbf{R}_{3a} + \mathbf{R}_{2a} \end{array}\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & 6 & -3 & 1 & 1 \end{array}\right]\begin{array}{l} \mathbf{R}_1 \\ \mathbf{R}_{2a} \\ \mathbf{R}_{3b} \end{array}$$

Now do the equivalent of back substitution.

$$\begin{array}{r} \\ \\ \mathbf{R}_{3b} \div 6 \end{array}\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{6} & \frac{1}{6} \end{array}\right]\begin{array}{l} \mathbf{R}_1 \\ \mathbf{R}_{2a} \\ \mathbf{R}_{3c} \end{array}$$

$$\begin{array}{r} \mathbf{R}_1 - 3\mathbf{R}_{3c} \\ \mathbf{R}_{2a} - 2\mathbf{R}_{3c} \\ \\ \end{array}\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & \frac{5}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & -1 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{6} & \frac{1}{6} \end{array}\right]\begin{array}{l} \mathbf{R}_{1a} \\ \mathbf{R}_{2b} \\ \mathbf{R}_{3c} \end{array}$$

$$\begin{array}{r} \mathbf{R}_{1a} - 2\mathbf{R}_{2b} \\ \\ \\ \end{array}\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{9}{2} & -\frac{11}{6} & \frac{1}{6} \\ 0 & 1 & 0 & -1 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{6} & \frac{1}{6} \end{array}\right]\begin{array}{l} \mathbf{R}_{1b} \\ \mathbf{R}_{2b} \\ \mathbf{R}_{3c} \end{array}$$

Finish with $\qquad \mathbf{I} \quad | \quad \mathbf{A}^{-1}$

(The way we have labelled the rows above keeps track of how they change. Thus the first time a row operation is done to $\mathbf{R}_2$, for example, it becomes $\mathbf{R}_{2a}$; the second time it becomes $\mathbf{R}_{2b}$. Once you have acquired some practice at doing row operations like this, you will probably find it unnecessary to keep labelling the rows.)

Exercise 4

Use Gaussian elimination-type operations to find the inverses of

(i) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ $\qquad$ (ii) $\begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 3 \\ 3 & 0 & 1 \end{bmatrix}$

[*Solution on p. 59*]

4.3 The existence of $\mathbf{A}^{-1}$

As you saw in Subsection 4.1, not all matrices have inverses. If we take a matrix which we know has no inverse, and attempt to use the method in Subsection 4.2, we shall find out how it breaks down.

Example 7

Suppose we attempt to find the inverse of

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

(for which $ad - bc \neq 0$). Using the method described in Subsection 4.2, we get

$$\begin{array}{rcl}
 & \mathbf{A} \qquad | \qquad \mathbf{I} & \\
 & \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{array}\right] & \begin{array}{l}\mathbf{R}_1\\ \mathbf{R}_2\end{array} \\
\begin{array}{r} \\ \mathbf{R}_2 - 2\mathbf{R}_1 \end{array} & \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{array}\right] & \begin{array}{l}\mathbf{R}_1\\ \mathbf{R}_{2a}\end{array}
\end{array}$$

At this point, we have to stop, as there is no multiple of $\mathbf{R}_2$ $[0 \quad 0]$ which we can subtract from $\mathbf{R}_1$ to reduce $\mathbf{A}$ to $\mathbf{I}$. In fact, if ever a row of zeros occurs in this way, we are not going to be able to complete the process to find the inverse. This means that the condition for $\mathbf{A}^{-1}$ to exist is precisely the same condition that

$$\mathbf{Ax} = \mathbf{b}$$

should have a unique solution—namely that the rows of $\mathbf{A}$ must be linearly independent.

You might get the feeling by now that everything seems to go wrong if the rows of a matrix are linearly dependent. This wouldn't be far from the truth! Since nearly every proof in linear algebra assumes that the rows of $\mathbf{A}$ are not linearly dependent, some terminology is needed to indicate briefly whether this is so.

> A **singular matrix** is a square matrix whose rows are linearly dependent. Conversely (and not surprisingly),
>
> a **non-singular matrix** is a square matrix whose rows are linearly independent.

> A matrix has an inverse if and only if it is non-singular.

As you saw in Subsection 4.2, the problem of finding an inverse is effectively the same as that of solving a set of simultaneous equations. Thus, the discussion about rounding error and ill-conditioning in the unit on simultaneous equations can be applied to the problem of finding an inverse. Partial pivoting is used in an effort to minimize induced ill-conditioning and the subsequent build-up of error. However, we are not going to enter further into these questions in this course.

Partial pivoting is discussed in Section 3 of *Unit 9*.

4.4 Matrix algebra

The laws we discussed in the earlier sections of this unit, and knowledge of specific matrices such as $\mathbf{0}$, $\mathbf{I}$ and $\mathbf{A}^{-1}$, enable us now to do some matrix algebra (as compared with matrix arithmetic). So long as we only use known facts about matrices we may manipulate the matrix blocks without any regard to their actual contents (unless we need to, of course!)

This subsection consists entirely of examples of matrix algebra, and some exercises for you to practise with.

Example 8

Multiplying out the expression $(\mathbf{A} + \mathbf{B})^2$, we get

$$\begin{aligned}
(\mathbf{A} + \mathbf{B})^2 &= (\mathbf{A} + \mathbf{B})(\mathbf{A} + \mathbf{B}) \\
&= \mathbf{A}(\mathbf{A} + \mathbf{B}) + \mathbf{B}(\mathbf{A} + \mathbf{B}) \qquad \text{(using the distributive law)} \\
&= \mathbf{A}^2 + \mathbf{AB} + \mathbf{BA} + \mathbf{B}^2.
\end{aligned}$$

It is important to note that this is as far as we can go, as in general $\mathbf{AB}$ does not equal $\mathbf{BA}$, so $\mathbf{AB} + \mathbf{BA}$ cannot be written as one term, as is the case for real numbers. However, there is no ambiguity in referring to $\mathbf{AA}$ as $\mathbf{A}^2$, etc.

Exercise 5

Expand the expression

$$(\mathbf{A} + \mathbf{B})(\mathbf{A} - \mathbf{B}).$$

[*Solution on p. 60*]

Example 9

We took as our definition of the inverse matrix that

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

Now we want to show that if $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, then $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$, which was stated without proof in Subsection 4.1. The argument can be set out as follows.

We assume that there is some matrix $\mathbf{Q}$ such that

$$\mathbf{A}\mathbf{Q} = \mathbf{I}. \tag{7}$$

We now want to show that $\mathbf{Q} = \mathbf{A}^{-1}$. To do this, we multiply both sides of (7) on the left by $\mathbf{A}^{-1}$.

$$\mathbf{A}^{-1}(\mathbf{A}\mathbf{Q}) = \mathbf{A}^{-1}\mathbf{I}, \text{ and so}$$

$$(\mathbf{A}^{-1}\mathbf{A})\mathbf{Q} = \mathbf{A}^{-1},$$

by the associative law and the properties of $\mathbf{I}$. That is,

$$\mathbf{Q} = \mathbf{A}^{-1}.$$

So we have proved that for any square matrix $\mathbf{A}$ which has an inverse $\mathbf{A}^{-1}$ such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, it is also true that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}.$$

Example 10

Show for non-singular matrices $\mathbf{A}$ and $\mathbf{B}$, that

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

This means that we have to show that

$$(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{A}\mathbf{B}) = \mathbf{I}.$$

Solution

$$\begin{aligned}
(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{A}\mathbf{B}) &= \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} && \text{(using the associative law)}\\
&= \mathbf{B}^{-1}\mathbf{I}\mathbf{B} && (\text{since } \mathbf{A}^{-1}\mathbf{A} = \mathbf{I})\\
&= \mathbf{B}^{-1}\mathbf{B} &&\\
&= \mathbf{I} && (\text{since } \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}).
\end{aligned}$$

So the inverse of $\mathbf{A}\mathbf{B}$ must be $\mathbf{B}^{-1}\mathbf{A}^{-1}$.

Exercise 6

Show for non-singular matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ that

$$(\mathbf{A}\mathbf{B}\mathbf{C})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}.$$

[*Solution on p. 60*]

It can be shown by induction that for non-singular matrices $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$,

$$\boxed{(\mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_n)^{-1} = \mathbf{A}_n^{-1}\mathbf{A}_{n-1}^{-1}\ldots\mathbf{A}_2^{-1}\mathbf{A}_1^{-1}.}$$

Exercise 7 (short!)

(i) What is the inverse of $\mathbf{A}^{-1}$?

(ii) What is the transpose of $\mathbf{A}^T$?

[*Solution on p. 60*]

The last example in this subsection concerns a type of matrix called a symmetric matrix.

A **symmetric matrix** $\mathbf{A}$ is defined to be one for which

$$\mathbf{A}^T = \mathbf{A}.$$

An example of a symmetric matrix is

$$\begin{bmatrix} 2 & 4 \\ 4 & 1 \end{bmatrix},$$

and, as you can see, it is symmetric about its main diagonal.

Example 11

Show that any matrix $\mathbf{Q}$ has the property that $\mathbf{QQ}^T$ is symmetric. (Note that the product $\mathbf{QQ}^T$ can always be formed, as the number of columns of $\mathbf{Q}$ is automatically equal to the number of rows of $\mathbf{Q}^T$.)

Solution

We need to show that $(\mathbf{QQ}^T)^T = \mathbf{QQ}^T$. Now, using the fact that $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ (see Subsection 2.5), we get

$$\begin{aligned}(\mathbf{QQ}^T)^T &= (\mathbf{Q}^T)^T\mathbf{Q}^T \\ &= \mathbf{QQ}^T, \qquad \text{since } (\mathbf{Q}^T)^T = \mathbf{Q} \text{ from Exercise 7(ii).}\end{aligned}$$

So the result is proved.

Exercise 8

Given that $\mathbf{A}^2 = \mathbf{I}$, show that

$$(\mathbf{I} + \mathbf{A})(\mathbf{I} - \mathbf{A}) = \mathbf{0}.$$

Warning: $\mathbf{A}^2 = \mathbf{I}$ *does not necessarily mean that* $\mathbf{A} = \pm\mathbf{I}$. *You can check for instance that* $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ *has this property.*

Exercise 9

Show that for any non-singular square matrix $\mathbf{A}$,

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T.$$

Exercise 10

(i) Show that if $\mathbf{A}$ is any square matrix, then $(\mathbf{A} + \mathbf{I})$ and $(\mathbf{A} - \mathbf{I})$ commute, i.e.

$$(\mathbf{A} + \mathbf{I})(\mathbf{A} - \mathbf{I}) = (\mathbf{A} - \mathbf{I})(\mathbf{A} + \mathbf{I}).$$

(ii) Use this to show that if $(\mathbf{A} + \mathbf{I})$ is a non-singular matrix, then $(\mathbf{A} + \mathbf{I})^{-1}$ and $(\mathbf{A} - \mathbf{I})$ commute, i.e.

$$(\mathbf{A} + \mathbf{I})^{-1}(\mathbf{A} - \mathbf{I}) = (\mathbf{A} - \mathbf{I})(\mathbf{A} + \mathbf{I})^{-1}.$$

[*Solutions to Exercises 8–10 on p. 60*]

Summary of Section 4

1. Definitions

(i) The **unit** (or **identity**) **matrix** $\mathbf{I}$ is a square matrix (whose size depends on the context), with ones down the main diagonal and zeros everywhere else. For any square matrix $\mathbf{A}$,

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}.$$

(ii) A **singular matrix** is a square matrix whose rows are linearly dependent. A **non-singular matrix** is a square matrix whose rows are linearly independent.

(iii) The **inverse matrix** $\mathbf{A}^{-1}$ of a non-singular matrix $\mathbf{A}$ has the property that

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{AA}^{-1} = \mathbf{I}.$$

Singular matrices (and non-square matrices) have no inverses.

(iv) A **symmetric matrix** $\mathbf{A}$ is one for which

$$\mathbf{A}^T = \mathbf{A}.$$

2. Computation of $\mathbf{A}^{-1}$

The problem of finding $\mathbf{A}^{-1}$ is effectively the same as solving the set of equations represented by

$$\mathbf{Ax} = \mathbf{b}.$$

We use Gaussian elimination type operations, with partial pivoting. Starting with the matrix pair

$$\mathbf{A}|\mathbf{I},$$

perform the same linear row operations simultaneously on both these matrices. If $\mathbf{A}$ is reduced to $\mathbf{I}$, then the same row operations will change $\mathbf{I}$ to $\mathbf{A}^{-1}$.

3. A property of inverses

$$(\mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_n)^{-1} = \mathbf{A}_n^{-1}\mathbf{A}_{n-1}^{-1}\ldots\mathbf{A}_2^{-1}\mathbf{A}_1^{-1}.$$

End of section exercises

Exercise 11

Use Gaussian elimination-type operations to find the inverses (where possible) of the following matrices:

(i) $\begin{bmatrix} 3 & 4 \\ 9 & 12 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & -2 \\ 3 & -1 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

Exercise 12

If the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{X}$ are non-singular and

$$\mathbf{B} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X},$$

find $\mathbf{A}$ in terms of $\mathbf{B}$ and $\mathbf{X}$.

Exercise 13

Show that if $\mathbf{A}$ and $\mathbf{B}$ are square matrices such that $\mathbf{A}$ and $\mathbf{B}$ commute, then:

(i) $\mathbf{A}^T$ and $\mathbf{B}^T$ commute,

(ii) $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ commute,

(iii) $\mathbf{A}$ and $\mathbf{B}^{-1}$ commute.

[*Solutions to Exercises 11–13 on pp. 60–61*]

5 Introduction to determinants

In the next few units, we shall be considering some applications of matrices—to vibrations of mechanical systems, for example. In addition to the matrix algebra considered so far in this unit, some of these applications involve a number associated with a square matrix, called its **determinant**.

5.1 Definition of 2×2 and 3×3 determinants

(a) 2×2 determinants

We have seen in Subsection 4.1 that the 2×2 matrix

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

has an inverse if and only if the number $ad - bc$ is not equal to zero. This number $ad - bc$ is called the **determinant** of the matrix $\mathbf{A}$, often denoted by $\det \mathbf{A}$. Instead of $\det \mathbf{A}$, we sometimes write

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix},$$

where the straight lines indicate that we are referring to the number $ad - bc$, *not* the matrix.

Example 1

If $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, then

$$\det \mathbf{A} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$$
$$= (1 \times 4) - (2 \times 3) = -2.$$

Exercise 1

Evaluate $\det \mathbf{A}$ if:

(i) $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ -4 & 5 \end{bmatrix}$ (ii) $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$.

[Solution on p. 61]

(b) The 3×3 case

A definition of the determinant of a 3×3 matrix can be given as follows. If

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

then

$$\begin{aligned} \det \mathbf{A} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \qquad (1) \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} \\ &\quad + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \end{aligned}$$

This way of evaluating $\det \mathbf{A}$ is often referred to as **expanding by the top row**. This is simply because we write down the first element in the top row (a_{11}) and multiply this by the determinant of the elements left when we cross out the row and the column in which a_{11} lies.

$$\begin{array}{ccc} \cancel{a_{11}} & \cancel{a_{12}} & \cancel{a_{13}} \\ \cancel{a_{21}} & a_{22} & a_{23} \\ \cancel{a_{31}} & a_{32} & a_{33} \end{array}$$

This is simply the determinant

$$\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}.$$

We then move to the next element in the top row (a_{12}), and multiply this by the determinant of the remaining elements when we cross out the row and column in which a_{12} lies.

$$\begin{array}{ccc} \cancel{a_{11}} & \cancel{a_{12}} & \cancel{a_{13}} \\ a_{21} & \cancel{a_{22}} & a_{23} \\ a_{31} & \cancel{a_{32}} & a_{33} \end{array}$$

This is the determinant

$$\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}.$$

We continue along the top row in this fashion until we reach the last element. As you can see from the definition, the signs connecting these terms alternate. A general $n \times n$ determinant may be defined in a similar way, but for the purposes of this section we will confine the discussion mostly to 2×2 and 3×3 determinants.

Example 2

If $\mathbf{A} = \begin{bmatrix} 2 & 1 & 3 \\ 0 & 2 & 1 \\ 3 & 1 & 6 \end{bmatrix}$,

then using Equation (1),

$$\begin{aligned}\det \mathbf{A} &= \begin{vmatrix} 2 & 1 & 3 \\ 0 & 2 & 1 \\ 3 & 1 & 6 \end{vmatrix} \\ &= \left(2 \times \begin{vmatrix} 2 & 1 \\ 1 & 6 \end{vmatrix}\right) - \left(1 \times \begin{vmatrix} 0 & 1 \\ 3 & 6 \end{vmatrix}\right) + \left(3 \times \begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix}\right) \\ &= (2 \times 11) - (-3) + (3 \times -6) \\ &= 22 + 3 - 18 \\ &= 7.\end{aligned}$$

Exercise 2

If $\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 2 & 3 \\ 4 & 1 & 5 \end{bmatrix}$,

evaluate (i) $\det \mathbf{A}$, (ii) $\det \mathbf{A}^T$.

Exercise 3

Find $\det \mathbf{A}$ if $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

[*Solutions to Exercises 2 and 3 on p. 61*]

You will see some of the situations in which determinants arise in the next subsection, and you will need to be able to evaluate determinants in the unit on eigenvalues.

5.2 Situations in which determinants arise

Surprisingly, determinants can be used to simplify the formulas occurring in various mathematical situations.

(a) Vector products

(*i*) *The cross product*

In *Unit 14* on vector algebra, you saw that the cross product $\mathbf{a} \times \mathbf{b}$ was given by the lengthy expression

$$\mathbf{a} \times \mathbf{b} = (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}.$$

We can re-express the middle term, to get

$$\begin{aligned}\mathbf{a} \times \mathbf{b} &= (a_2b_3 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} \\ &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\mathbf{k}. \qquad (2)\end{aligned}$$

This expression can be remembered even more conveniently as a 3×3 determinant:

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}. \qquad (3)$$

Note that we take the liberty of putting $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ in the top row, even though they aren't numbers. This is because the resulting form is so easy to memorize.

Exercise 4

Use the determinant form of the cross product to evaluate $\mathbf{a} \times \mathbf{b}$, where

$$\mathbf{a} = 3\mathbf{i} - \mathbf{j} + 4\mathbf{k} \quad \text{and} \quad \mathbf{b} = \mathbf{i} - \mathbf{j} + \mathbf{k}.$$

[*Solution on p. 61*]

(ii) The triple scalar product

The triple scalar product $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$ which you met in *Unit 14* can also be most easily remembered as a 3×3 determinant.

Expressing $\mathbf{b}\times\mathbf{c}$ in the style of (2) above:

$$\mathbf{b}\times\mathbf{c} = \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}\mathbf{k},$$

and the dot product of $a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ with this is then

$$\begin{aligned}\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) &= a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} \\ &= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \qquad (4)\end{aligned}$$

Again, this is far the simplest way of writing the triple scalar product.

(b) Areas and volumes

(i) The area of a triangle in two dimensions

The area of a triangle in two dimensions can be expressed as a 3×3 determinant. From *Unit 14* you know that

$$\overrightarrow{AB} \times \overrightarrow{AC} = (|\overrightarrow{AB}||\overrightarrow{AC}|\sin\theta)\mathbf{k}$$

where the triangle ABC is as in Figure 1, θ is the angle between $\overrightarrow{AB}$ and $\overrightarrow{AC}$ (both in the xy-plane), and $\mathbf{k}$ is the Cartesian unit vector perpendicular to the xy-plane. We can also express the area of the triangle ABC as

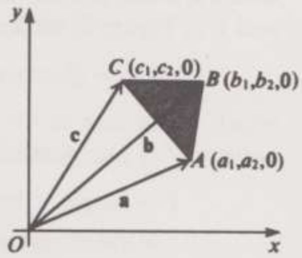


Figure 1

$$\begin{aligned}\text{Area of triangle } ABC &= \tfrac{1}{2}|\overrightarrow{AB}||\overrightarrow{AC}|\sin\theta \\ &= \tfrac{1}{2}|\overrightarrow{AB}\times\overrightarrow{AC}|. \qquad (5)\end{aligned}$$

Now,

$$\begin{aligned}\overrightarrow{AB}\times\overrightarrow{AC} &= (\mathbf{b}-\mathbf{a})\times(\mathbf{c}-\mathbf{a}) \\ &= (\mathbf{b}\times\mathbf{c}) - (\mathbf{a}\times\mathbf{c}) - (\mathbf{b}\times\mathbf{a}) + (\mathbf{a}\times\mathbf{a}) \\ &= (\mathbf{b}\times\mathbf{c}) - (\mathbf{a}\times\mathbf{c}) + (\mathbf{a}\times\mathbf{b}) \\ &\qquad\qquad \text{(by the properties of cross products)} \\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ b_1 & b_2 & 0 \\ c_1 & c_2 & 0 \end{vmatrix} - \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & 0 \\ c_1 & c_2 & 0 \end{vmatrix} + \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & 0 \\ b_1 & b_2 & 0 \end{vmatrix} \\ &= ((b_1c_2 - c_1b_2) - (a_1c_2 - c_1a_2) + (a_1b_2 - b_1a_2))\mathbf{k}.\end{aligned}$$

The coefficient of $\mathbf{k}$ in this determinant can be written as the determinant

$$= \begin{vmatrix} 1 & 1 & 1 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix},$$

and so from Equation (5),

$$\text{Area of triangle } ABC = \text{modulus of } \tfrac{1}{2}\begin{vmatrix} 1 & 1 & 1 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix}. \qquad (6)$$

(We have had to write 'modulus of' because of the unfortunate similarity between the modulus sign and the above way of writing determinants.)

Exercise 5

Show that the area of the triangle OAB in Figure 1 is

$$\tfrac{1}{2}\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

[*Solution on p. 61*]

(ii) Volumes in three dimensions

The volume of the squashed box in Figure 2 (normally called a **parallelepiped**) can be expressed as a determinant. Suppose the position vectors $\overrightarrow{OA}$, $\overrightarrow{OB}$ and $\overrightarrow{OC}$ are

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k},$$
$$\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k},$$
$$\mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$$

respectively, and $\mathbf{n}$ is a unit vector perpendicular to the base $OBDC$ of the parallelepiped.

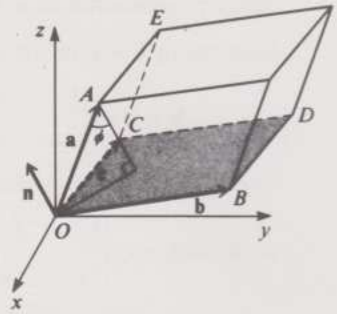


Figure 2

We know that the height h of the parallelepiped can be expressed as

$$\begin{aligned} h &= |\mathbf{a}|\cos\phi \\ &= |\mathbf{a}\cdot\mathbf{n}|, \end{aligned} \tag{7}$$

where $\mathbf{a}$, $\mathbf{n}$ and ϕ are as shown in Figure 2. Also, if θ is the angle between the vectors $\mathbf{b}$ and $\mathbf{c}$ in Figure 2, we know from *Unit 14* that

$$\mathbf{b}\times\mathbf{c} = (|\mathbf{b}||\mathbf{c}|\sin\theta)\mathbf{n}, \tag{8}$$

which is the area of the base of the parallelepiped, times a unit vector $\mathbf{n}$. Now, the volume V of our parallelepiped is

$$\begin{aligned} V &= \text{height} \times \text{base} \\ &= (|\mathbf{a}|\cos\phi)\times(|\mathbf{b}||\mathbf{c}|\sin\theta) \\ &= |\mathbf{a}\cdot\mathbf{n}|\times(|\mathbf{b}||\mathbf{c}|\sin\theta) && \text{from (7)} \\ &= |\mathbf{a}\cdot(|\mathbf{b}||\mathbf{c}|\sin\theta\mathbf{n})| \\ &= |\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})| && \text{from (8).} \end{aligned}$$

Thus, using Equation (4), we derive a determinant expression for the volume:

$$V = \text{modulus of} \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \tag{9}$$

Exercise 6

(i) Find the volume of the parallelepiped generated (as in Figure 2) by the position vectors

$$\mathbf{a} = \mathbf{i} + \mathbf{k}, \qquad \mathbf{b} = \mathbf{i} + 2\mathbf{j} \qquad \text{and} \qquad \mathbf{c} = \mathbf{j} + 3\mathbf{k}.$$

(ii) Show for this particular case that

$$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}) = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}).$$

[*Solution on p. 61*]

You can see from Exercise 6 that cyclic permutations of $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ gave the same numerical result. This is fairly obvious if we think about the way we generated the volume of the parallelepiped. The base was quite arbitrary. We could have just as easily chosen $OCEA$ as the base, and its area would be the magnitude of $\mathbf{c}\times\mathbf{a}$. So long as we are careful not to alter the right-handed order of the vectors, any of the expressions in Exercise 6 (ii) will generate the volume of the figure.

(c) Solutions of simultaneous equations

You can check that the equations

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$

have the solution

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{22}a_{11} - a_{21}a_{12}}, \qquad x_2 = \frac{a_{21}b_1 - a_{11}b_2}{a_{21}a_{12} - a_{11}a_{22}}$$

provided that the denominators aren't zero. Writing the equations in matrix form $\mathbf{Ax} = \mathbf{b}$, the solution can be expressed in determinant form

$$x_1 = \frac{\Delta_1}{\Delta}, x_2 = \frac{\Delta_2}{\Delta}$$

where Δ is $\det \mathbf{A}$, Δ_1 is the determinant of the matrix formed by replacing the first column of $\mathbf{A}$ by $\mathbf{b}$, and Δ_2 is the determinant formed in a similar way by replacing the second column of $\mathbf{A}$ by $\mathbf{b}$.

There is a similar formula for the solution of any system $\mathbf{Ax} = \mathbf{b}$ of n equations in n unknowns, known as **Cramer's rule**:

$$x_1 = \frac{\Delta_1}{\Delta}, x_2 = \frac{\Delta_2}{\Delta}, \ldots, x_n = \frac{\Delta_n}{\Delta} \tag{10}$$

provided that $\Delta \neq 0$. Here, Δ is $\det \mathbf{A}$, and Δ_k is the determinant of the matrix formed by replacing the kth column of $\mathbf{A}$ by $\mathbf{b}$.

As you will see in Subsection 5.4, it turns out in practice that Gaussian elimination is a much more efficient way of solving simultaneous equations. However, Cramer's rule provides an interesting theoretical view. For example, expressing the solutions in the form of Equation (10), we can rephrase in terms of determinants the conditions which we found in *Unit 9* for a unique solution, no solution and an infinite number of solutions.

1. If $\Delta \neq 0$, the equations have a unique solution.
2. If $\Delta = 0$ and one or more of the Δ_ks is non-zero, then the equations are inconsistent, and have no solution.
3. If $\Delta = 0$ and also $\Delta_1 = \Delta_2 = \cdots = \Delta_n = 0$, then the equations are linearly dependent, and will usually have an infinite number of solutions.

5.3 Properties of determinants

Determinants have some interesting and useful properties. In this subsection we explore some of these for 3×3 determinants, using Equation (4) which tells us that

$$\Delta = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$$

and using some of the properties of vectors which you have already met in *Unit 14*, and Subsection 5.2 of this unit. In particular, we shall be using the cyclic property which you met in Exercise 6:

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}). \tag{11}$$

Property 1: *interchanging rows multiplies the determinant by* -1.

To examine the effect of interchanging the second and third row, we look at $\mathbf{a} \cdot (\mathbf{c} \times \mathbf{b})$. Now we know that

$$\mathbf{c} \times \mathbf{b} = -\mathbf{b} \times \mathbf{c},$$

so

$$\begin{aligned} \mathbf{a} \cdot (\mathbf{c} \times \mathbf{b}) &= -\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) \\ &= -\Delta. \end{aligned} \tag{12}$$

If we change the first two rows, we get

$$\begin{aligned} \mathbf{b} \cdot (\mathbf{a} \times \mathbf{c}) &= -\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) \\ &= -\Delta \qquad \text{(from (11))}. \end{aligned}$$

The following exercise completes the possibilities.

Exercise 7

Show that $\mathbf{c}\cdot(\mathbf{b}\times\mathbf{a}) = -\Delta$.

[*Solution on p. 62*]

So, interchanging any two rows of a determinant changes the sign of the determinant. We shall refer briefly to this property as P1, and to the following properties as P2, P3, etc.

Property 2: *if two rows of a determinant Δ are identical, then $\Delta = 0$.*

If row 2 and row 3 are identical, then

$$\begin{aligned}\Delta &= \mathbf{a}\cdot(\mathbf{b}\times\mathbf{b})\\ &= 0 \qquad \text{(a property of the cross product).}\end{aligned}$$

If row 1 and row 2 are identical, then

$$\begin{aligned}\Delta &= \mathbf{a}\cdot(\mathbf{a}\times\mathbf{c})\\ &= \mathbf{c}\cdot(\mathbf{a}\times\mathbf{a}) \qquad \text{(from (11))}\\ &= 0.\end{aligned}$$

Exercise 8

Show that if row 1 and row 3 are identical, then $\Delta = 0$.

[*Solution on p. 62*]

Property 3: *if a row of a determinant Δ is multiplied by k, then the value of this determinant is $k\Delta$.*

We know that multiplying either of the vectors in a dot or cross product by k multiplies the result by k. Thus,

$$\begin{aligned}(k\mathbf{a})\cdot(\mathbf{b}\times\mathbf{c}) &= k(\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}))\\ &= k\Delta.\end{aligned}$$

Also, $\mathbf{a}\cdot((k\mathbf{b})\times\mathbf{c})$ and $\mathbf{a}\cdot(\mathbf{b}\times(k\mathbf{c}))$ are each equal to $\mathbf{a}\cdot(k(\mathbf{b}\times\mathbf{c}))$, which is equal to $k\Delta$.

This result isn't very important in its own right, but it helps to prove the next property, which is very useful.

Property 4: *if one row of Δ is a linear combination of the others, then $\Delta = 0$.*

Suppose row 1 is a linear combination $k_1\mathbf{b} + k_2\mathbf{c}$ of row 2 and row 3. Then

$$\begin{aligned}\Delta &= (k_1\mathbf{b} + k_2\mathbf{c})\cdot(\mathbf{b}\times\mathbf{c})\\ &= (k_1\mathbf{b})\cdot(\mathbf{b}\times\mathbf{c}) + (k_2\mathbf{c})\cdot(\mathbf{b}\times\mathbf{c})\\ &= k_1(\mathbf{b}\cdot(\mathbf{b}\times\mathbf{c})) + k_2(\mathbf{c}\cdot(\mathbf{b}\times\mathbf{c}))\\ &= 0 \qquad \text{(from P2).}\end{aligned}$$

Similarly, if row 2 is a linear combination $k_1\mathbf{a} + k_2\mathbf{c}$ of row 1 and row 3, then

$$\begin{aligned}\Delta &= \mathbf{a}\cdot((k_1\mathbf{a} + k_2\mathbf{c})\times\mathbf{c})\\ &= \mathbf{a}\cdot((k_1\mathbf{a})\times\mathbf{c}) + \mathbf{a}\cdot((k_2\mathbf{c})\times\mathbf{c})\\ &= k_1(\mathbf{a}\cdot(\mathbf{a}\times\mathbf{c})) + k_2(\mathbf{a}\cdot(\mathbf{c}\times\mathbf{c})) \qquad \text{(from P3)}\\ &= 0 \qquad \text{(from P2).}\end{aligned}$$

Exercise 9

Show that if row 3 is a linear combination of rows 1 and 2, then $\Delta = 0$.

[*Solution on p. 62*]

The consequences of P4 are wide and will be discussed in Subsection 5.4.

Property 5: *if a row of Δ is zero, then $\Delta = 0$.*

This is just a statement of P4, if one row were zero times another one.

Property 6: *if a linear combination of a set of rows of a determinant is added to another row not in the set, then the value of the determinant remains unchanged.*

Suppose we add k_1 times row 1 plus k_2 times row 2 to row 3, for instance. Then

$$\begin{aligned}\mathbf{a}\cdot(\mathbf{b}\times(\mathbf{c}+k_1\mathbf{a}+k_2\mathbf{b})) &= \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})+\mathbf{a}\cdot(\mathbf{b}\times(k_1\mathbf{a}+k_2\mathbf{b}))\\ &= \Delta+0 \qquad \text{(from P4).}\end{aligned}$$

The other possibilities follow in the same way.

The last property we discuss in any detail in this section will be proved algebraically without using vectors.

Property 7: $\det \mathbf{A} = \det \mathbf{A}^T$.

Proof

$$\begin{aligned}\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} &= a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}\\ &= a_1(b_2c_3-b_3c_2)-a_2(b_1c_3-b_3c_1)+a_3(b_1c_2-b_2c_1)\\ &= a_1(b_2c_3-c_2b_3)-b_1(a_2c_3-c_2a_3)+c_1(a_2b_3-b_2a_3)\\ &= a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}\\ &= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \det \mathbf{A}^T.\end{aligned}$$

Property 7 has an important consequence. Properties 1 to 6 can be applied to any determinant. In particular, these properties can be used to describe the effect which ROW operations on $\mathbf{A}^T$ have on the value of $\det \mathbf{A}^T$ (and hence the value of $\det \mathbf{A}$, using Property 7). Now *row* operations on $\mathbf{A}^T$ are *column* operations on $\mathbf{A}$. The consequence of this is that in Properties 1 to 6, the word 'row' could be replaced by 'column', and the statements would still be true. For example, rewriting Property 5: if a column of Δ is zero, then $\Delta = 0$.

The following example demonstrates how to use these properties to simplify the evaluation of determinants.

Example 3

Evaluate $\det \mathbf{A}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

(i) Using P6 twice, we can subtract 4 times row 1 from row 2 and then 7 times row 1 from row 3, to obtain

$$\det \mathbf{A} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{vmatrix}.$$

Again using P6, we can subtract twice the new row 2 from the new row 3 to obtain

$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{vmatrix} = 0 \qquad \text{(from P5).}$$

(ii) Another way of evaluating this determinant would have used column operations. We can use P6 twice, to subtract twice column 1 from column 2 and then 3 times column 1 from column 3, to obtain

$$\det \mathbf{A} = \begin{vmatrix} 1 & 0 & 0 \\ 4 & -3 & -6 \\ 7 & -6 & -12 \end{vmatrix}.$$

Now, subtracting twice the new column 2 from the new column 3, we obtain

$$\det \mathbf{A} = \begin{vmatrix} 1 & 0 & 0 \\ 4 & -3 & 0 \\ 7 & -6 & 0 \end{vmatrix} = 0 \qquad \text{(from P5).}$$

There are plenty of other ways of showing that this determinant is zero, but the Gaussian elimination technique, which P6 allows us to do, is as good as any.

A look back to P4 tells us that if the rows of **A** are linearly dependent then the value of the determinant will be zero. The four statements about a square matrix:

(i) the rows of **A** are linearly dependent,

(ii) **A** is a singular matrix,

(iii) $\det \mathbf{A} = 0$,

(iv) $\mathbf{A}^{-1}$ does not exist,

all say the same thing in different ways. Equally well, the statements:

(i) the rows of **A** are linearly independent,

(ii) **A** is a non-singular matrix,

(iii) $\det \mathbf{A} \neq 0$,

(iv) $\mathbf{A}^{-1}$ exists,

all say the same thing. We shall find in the next unit how the determinant form of this statement is useful.

There is one other useful property which, for completeness, is stated in this section, although not proved.

Property 8: *given two square matrices* **A** *and* **B** *of the same size, then*

$$\det(\mathbf{AB}) = \det \mathbf{A} \det \mathbf{B}.$$

Finally, although the proofs in this section only show that the properties are true for 3×3 determinants, they are in fact true for determinants of any size.

5.4 A general method for the evaluation of determinants

Any $n \times n$ determinant may be evaluated by an extension of the way we evaluated a 3×3 determinant in Subsection 5.1. However, in general it is much simpler to use a procedure analogous to Gaussian elimination. The idea of the method is to reduce the matrix to upper triangular form, since the determinant of an upper triangular matrix is very easy to evaluate. This is illustrated in the following example.

Example 4

To evaluate the determinant of an upper triangular matrix **A**, where

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}.$$

P7 tells us (among other things) that expanding by columns is just as valid as expanding by rows, so we expand by the first column to obtain

$$\det \mathbf{A} = a\begin{vmatrix} d & e \\ 0 & f \end{vmatrix} - 0\begin{vmatrix} b & c \\ 0 & f \end{vmatrix} + 0\begin{vmatrix} b & c \\ d & e \end{vmatrix}$$
$$= adf.$$

It can be shown that this principle generalizes to the case of an $n \times n$ matrix in upper triangular form. We thus have the following result.

> If any square matrix is in **upper triangular form** (i.e. all elements below the main diagonal are zeros), then its determinant is just the product of its diagonal elements.

Now, we know from P6 that the row operations performed in Gaussian elimination without pivoting do not affect the value of the determinant. (Remember, the only operation allowed us in Gaussian elimination without pivoting is to add or subtract a multiple of one row to or from another.) So, to find det **A**, use the operations of Gaussian elimination to reduce **A** to a matrix **U** in upper triangular form,

$$\mathbf{U} = \begin{vmatrix} u_{11} & \cdots\cdots & u_{1n} \\ & u_{22} \quad \cdots & u_{2n} \\ 0 & \ddots & u_{nn} \end{vmatrix}$$

Then

$$\det \mathbf{A} = \det \mathbf{U} = u_{11}u_{22}\ldots u_{nn}.$$

Example 5

Use this technique to evaluate det **A**, where

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 3 \\ 2 & 3 & 8 \end{bmatrix}.$$

First we do the Gaussian elimination steps

$$\mathbf{R}_{2a} = \mathbf{R}_2 - \mathbf{R}_1 \qquad \text{and} \qquad \mathbf{R}_{3a} = \mathbf{R}_3 - 2\mathbf{R}_1.$$

We know from P6 that this will not affect the value of the determinant, so

$$\det \mathbf{A} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 1 & 6 \end{vmatrix} \begin{matrix} \mathbf{R}_1 \\ \mathbf{R}_{2a} \\ \mathbf{R}_{3a} \end{matrix} \quad \text{(Stage 1(a))}$$

Subtracting $\frac{1}{3}\mathbf{R}_{2a}$ from $\mathbf{R}_{3a}$ gives

$$\det \mathbf{A} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & \frac{16}{3} \end{vmatrix} \qquad \text{(Stage 1(b))}.$$

At this stage, we can write down the value of det **A**:

$$\begin{aligned} \det \mathbf{A} &= 1 \times 3 \times \tfrac{16}{3} \\ &= 16. \end{aligned}$$

For large determinants the amount of work involved in evaluating det **A** using Gaussian elimination is far less than without it.

Exercise 10

If

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 7 \\ 3 & 4 & 1 \\ -1 & 2 & 7 \end{bmatrix},$$

find det **A** (i) directly, (ii) using Gaussian elimination.

Exercise 11

Use Gaussian elimination to evaluate $\begin{vmatrix} 0 & 2 & 3 \\ 1 & 3 & 6 \\ 2 & 2 & 4 \end{vmatrix}$.

[*Solutions to Exercises 10 and 11 on p. 62*]

Finally, let us take a brief look at Cramer's method for solving n simultaneous equations in n unknowns,

$$x_1 = \frac{\Delta_1}{\Delta}, \qquad x_2 = \frac{\Delta_2}{\Delta}, \qquad \ldots, \qquad x_n = \frac{\Delta_n}{\Delta}$$

where the Δ_ks are the determinants of the form given in Subsection 5.2. To solve a set of equations in this way would involve the equivalent of $n+1$ lots of Gaussian

elimination, whereas solving the equations directly only involves using Gaussian elimination once. So, although Cramer's rule is a neat theoretical solution, in practice we would hurriedly discard it in favour of Gaussian elimination.

Summary of Section 5

A **determinant**, denoted by $\det \mathbf{A}$, is a number associated with a given square matrix $\mathbf{A}$. If $\mathbf{A}$ is a 2×2 matrix, then

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

If $\mathbf{A}$ is a 3×3 matrix, then

$$\begin{aligned} \det \mathbf{A} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \end{aligned}$$

Larger determinants have similar definitions.

The most important properties of determinants are:

(i) if the rows (or columns) of $\mathbf{A}$ are linearly dependent, then $\det \mathbf{A} = 0$,

(ii) linear combinations of rows (or columns) of a determinant can be added to any other row (or column) without changing its value.

Using this last property, an $n \times n$ determinant $\det \mathbf{A}$ can be evaluated by reducing $\mathbf{A}$ to an **upper triangular matrix U**, using the operations of Gaussian elimination. Then

$$\det \mathbf{A} = \det \mathbf{U} = u_{11}u_{22}\ldots u_{nn}.$$

End of section exercise

Exercise 12

Find $\det \mathbf{A}$ if:

(i) $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$

(ii) $\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 1 & 2 \\ 2 & -1 & 1 \end{bmatrix}$

(iii) $\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 1 & 2 & 4 & -1 \\ 2 & 4 & -1 & 1 \\ 0 & 1 & 2 & -1 \end{bmatrix}$.

[*Solution on p. 62*]

6 End of unit test

As a lot of the terminology in this unit may be new to you, you will find a glossary of new terms used in Sections 1–5, and some important results, immediately following the test.

The test is divided into two sections. Section A should not take you longer than 10 minutes. If you get any of this section wrong, and don't understand why, go back and check the appropriate section in the unit before you go on to Section B.

Section A

There is only one correct option to each of the following questions.

1. Which of the following is a 3×2 matrix?

(a) $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 0 & 3 \\ 3 & 0 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 0.2 & 1.2 \\ 0.3 & 1.3 \\ 0.4 & 1.4 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & 3 & 1 \\ -1 & 4 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

(e) $[a \quad b \quad c \quad d \quad e \quad f]$

2. Given that

$$\begin{bmatrix} 8 & 5 \\ 2 & 1 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} x & r \\ y & s \\ z & t \end{bmatrix},$$

the value of s is

(a) $2 + 1$ (b) $5 + 1 + 4$ (c) $8r + 5t$ (d) 1 (e) $8x + 2y + 3z$.

3. If $\mathbf{A} = \begin{bmatrix} 4 & 3 & 5 \\ 2 & 1 & 6 \end{bmatrix}$, which of the following can be added to $\mathbf{A}$?

(a) $\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}$ (c) $[5 \quad 3 \quad 5]$ (d) $\begin{bmatrix} 5 & 3 & 5 \\ 3 & 1 & 5 \end{bmatrix}$ (e) $\begin{bmatrix} 1 \\ 5 \end{bmatrix}$

4. $\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} -1 & 1 \\ 2 & 0 \end{bmatrix}$ is equal to

(a) $\begin{bmatrix} 0 & 2 \\ 5 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$ (c) $[6 \quad 6]$ (d) $\begin{bmatrix} 3 & 0 \\ 3 & 6 \end{bmatrix}$ (e) $\begin{bmatrix} 2 \\ 9 \end{bmatrix}$

5. $2\begin{bmatrix} 8 & 12 & 4 \\ 4 & 30 & 2 \\ 2 & 10 & 0 \end{bmatrix}$ is equal to

(a) $\begin{bmatrix} 16 & 24 & 8 \\ 8 & 60 & 4 \\ 4 & 20 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 16 & 24 & 8 \\ 4 & 30 & 2 \\ 2 & 10 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 16 & 12 & 4 \\ 8 & 30 & 2 \\ 4 & 10 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} 4 & 6 & 2 \\ 2 & 15 & 1 \\ 1 & 5 & 0 \end{bmatrix}$ (e) $\begin{bmatrix} 4 & 6 & 2 \\ 4 & 30 & 2 \\ 2 & 10 & 0 \end{bmatrix}$

6. If $\mathbf{x} = [x_1 \quad x_2 \quad x_3]$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, then $\mathbf{xy}$ equals

(a) $[x_1y_1 + x_2y_2 + x_3y_3]$ (b) $\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix}$ (c) $\begin{bmatrix} x_1y_1 \\ x_2y_2 \\ x_3y_3 \end{bmatrix}$

(d) $[x_1y_1 \quad x_2y_2 \quad x_3y_3]$ (e) $\begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}$

7. Which one of the following can be multiplied on the left by $\begin{bmatrix} 12 & 13 & 4 \\ 4 & 56 & -50 \end{bmatrix}$?

(a) $\begin{bmatrix} 2 & 1 & 5 \\ 2 & 5 & 4 \end{bmatrix}$ (b) $\begin{bmatrix} 4 & 1 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} 4 & 3 & 6 & 8 \\ 1 & 1 & 4 & 2 \\ 3 & 5 & 3 & 5 \end{bmatrix}$

(d) $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ (e) $\begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}$

8. $\begin{bmatrix} 2 & 1 & 3 \\ 3 & 0 & 4 \end{bmatrix}\begin{bmatrix} 3 & 1 & 1 \\ 1 & 5 & 0 \\ 4 & -2 & 0 \end{bmatrix}$ is equal to

(a) $\begin{bmatrix} 19 & 0 & 4 \\ 2 & 3 & 1 \\ 25 & -5 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 6 & 5 & 4 \\ 18 & -3 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 19 & 1 & 2 \\ 25 & -5 & 3 \end{bmatrix}$

(d) $\begin{bmatrix} 19 & 25 \\ 1 & -5 \\ 2 & 3 \end{bmatrix}$ (e) $\begin{bmatrix} 6 & 1 & 3 \\ 3 & 0 & 0 \\ 8 & -2 & 0 \end{bmatrix}$

9. A matrix equivalent of the equations

$$\begin{aligned} 3x_1 + 2x_2 + x_3 &= 4 \\ x_2 &= x_3 \\ x_1 + 2x_2 &= x_3 \end{aligned}$$

is

(a) $\begin{bmatrix} 3 & 2 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$ (b) $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 3 & 2 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$

(c) $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\begin{bmatrix} 3 & 2 & 1 \\ 0 & 1 & -1 \\ 1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix}$ (d) $\begin{bmatrix} 3 & 2 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

(e) $\begin{bmatrix} 3 & 2 & 1 \\ 0 & 1 & -1 \\ 1 & 2 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix}$

10. If $\mathbf{Ax} = \mathbf{b}$, and $\mathbf{A}$ is non-singular, then

(a) $\mathbf{x} = \mathbf{bA}^{-1}$ (b) $\mathbf{bx} = \mathbf{A}^{-1}$ (c) $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ (d) $\mathbf{x} = \mathbf{IA}$ (e) $\mathbf{x} = \mathbf{AbA}^{-1}$

[*Solutions to Section A on p. 62*]

Section B

1. If $\mathbf{X} = \begin{bmatrix} 3 & 1 & 3 \\ 1 & 0 & 0 \\ 0 & 2 & 3 \end{bmatrix}$, evaluate

(i) $\mathbf{X} + \mathbf{I}$ (ii) $\mathbf{X}^2 + \mathbf{X}$ (iii) $\mathbf{X}^3 + \mathbf{X}^2$.

2. If $\mathbf{A}$ is a square matrix such that $\mathbf{A}^2 = \mathbf{0}$, show that

$$\mathbf{A}(\mathbf{I} + \mathbf{A})^2 = \mathbf{A}.$$

3. If $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are $n \times n$ matrices:

(i) simplify $(\mathbf{A} + \mathbf{B}) - (\mathbf{A} + \mathbf{C})$,

(ii) show that $(\mathbf{A} + \mathbf{B})^2 - \mathbf{A}(\mathbf{A} + 3\mathbf{B}) - \mathbf{B}(\mathbf{A} - \mathbf{B}) = 2(\mathbf{B} - \mathbf{A})\mathbf{B}$.

4.

(i) Find the inverse of the matrix

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ x & 1 & 0 \\ y & z & 1 \end{bmatrix}.$$

(ii) Given that

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 1 \\ 0 & 3 & 7 \\ 1 & 0 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 4 & 7 \\ 2 & 1 & 3 \end{bmatrix} \text{ and } \mathbf{c} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix},$$

find the column vector $\mathbf{x}$ with the property that

$$\mathbf{Bx} = \mathbf{Ax} + \mathbf{c}.$$

5.
(i) If $\mathbf{A}$ is a square matrix such that

$$\mathbf{A}^2 - \mathbf{A} + \mathbf{I} = \mathbf{0},$$

show that

$$\mathbf{A}^{-1} = \mathbf{I} - \mathbf{A}.$$

(ii) Show that

$$(\mathbf{R}^{-1}\mathbf{AR})^3 = \mathbf{R}^{-1}\mathbf{A}^3\mathbf{R}$$

where $\mathbf{R}$ is any non-singular matrix of the same size as $\mathbf{A}$. Generalize this result.

6. A square matrix $\mathbf{A}$ is said to be **orthogonal** if $\mathbf{AA}^T = \mathbf{I}$.

(i) Show that $\mathbf{A} = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}$ is orthogonal.

(ii) Write down the inverse of this matrix.

(iii) Show that if $\mathbf{A}$ and $\mathbf{B}$ are orthogonal matrices of the same size, then $\mathbf{AB}$ is also orthogonal.

7. Suppose $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is an $n \times 1$ column vector whose elements are the n variables $x_1, \dots, x_n$, and $\mathbf{A}$ is an $n \times n$ matrix.

Then the matrix

$$\mathbf{Q} = \mathbf{x}^T\mathbf{Ax},$$

regarded as a function of the variables $x_1, \dots, x_n$, is called a **quadratic form**.

(i) What is the size of $\mathbf{Q}$?

(ii) Find $\mathbf{Q}$ when

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{and} \quad \mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

(iii) Find a matrix $\mathbf{B}$ such that $\mathbf{x}^T\mathbf{Bx}$ is the quadratic form

$$x_1^2 + 2x_1x_2 + x_2^2.$$

(Note: there are several possible solutions to this part.)

8. (More difficult)

Suppose $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ are column vectors ($m \times 1$ and $n \times 1$ respectively), whose elements are variables, as in Question 7, and that $\mathbf{A}$ is an $n \times m$ matrix. Show that:

(i) $\mathbf{y}^T\mathbf{Ax} = (\mathbf{y}^T\mathbf{Ax})^T$,

(ii) $(\mathbf{y} - \mathbf{Ax})^T(\mathbf{y} - \mathbf{Ax}) = \mathbf{y}^T\mathbf{y} - 2\mathbf{y}^T\mathbf{Ax} + \mathbf{x}^T\mathbf{A}^T\mathbf{Ax}$.

9. Evaluate the determinants of the following matrices.

(i) $\begin{bmatrix} 4 & 1 \\ -2 & 8 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 0 & -3 \\ -1 & 1 & 0 \\ 2 & 3 & -1 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 3 & 0 & 4 \\ 0 & -1 & 2 & 1 \\ 3 & 0 & 1 & -1 \\ -2 & 3 & 1 & 0 \end{bmatrix}$

[*Solutions to Section B on pp. 62–64*]

Glossary of new terms used in this unit

1. *Matrix:* a rectangular block of numbers.
2. *Matrices*: plural of matrix.
3. *Column vector*: a matrix with only one column.
4. *An $m \times n$ matrix*: a matrix with m rows and n columns.
5. *Element*: an individual entry in a matrix.
6. a_{ij}: the element in the ith row and jth column of the matrix $\mathbf{A}$.
7. *Linear transformation*: a relationship between two column vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ that can be written in the form $\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1$.
8. $\mathbf{0}$, *a zero matrix*: a matrix each of whose elements is zero.
9. $\mathbf{I}$, *a unit matrix*: a square matrix with ones down the main diagonal, and zeros everywhere else.
10. $\mathbf{A}^{-1}$, *the inverse of* $\mathbf{A}$: a matrix with the property $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.
11. $\mathbf{A}^T$, *the transpose of* $\mathbf{A}$: a matrix whose columns are the rows of $\mathbf{A}$.
12. *A singular matrix*: a square matrix whose rows are linearly dependent.
13. *A non-singular matrix*: a square matrix whose rows are linearly independent.
14. *A symmetric matrix*: a matrix $\mathbf{A}$ with the property $\mathbf{A}^T = \mathbf{A}$.
15. *Upper triangular form*: refers to a square matrix with zeros below the main diagonal.
16. *An orthogonal matrix*: a square matrix $\mathbf{A}$ with the property that $\mathbf{A}\mathbf{A}^T = \mathbf{I}$.
17. *A quadratic form*: a single element matrix of the form $\mathbf{x}^T\mathbf{A}\mathbf{x}$.

Some matrix results

1. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.
2. $(\mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_n)^T = \mathbf{A}_n^T\mathbf{A}_{n-1}^T\ldots\mathbf{A}_2^T\mathbf{A}_1^T$.
3. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
4. $(\mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_n)^{-1} = \mathbf{A}_n^{-1}\mathbf{A}_{n-1}^{-1}\ldots\mathbf{A}_2^{-1}\mathbf{A}_1^{-1}$.

Appendix

Solutions to the exercises in Section 1

1. **(i)** Any matrix with two rows and two columns will do—for example, $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$.

(ii) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 1 & 2 & 3 \end{bmatrix}$ is a typical example.

(iii) $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$. This is an example of a column vector (or a column matrix).

(iv) $[1 \quad 4]$. This is an example of a row vector (or a row matrix).

2. This element would be referred to as p_{45}. (The row is always specified before the column.)

3. By the definition of equality we get

(i) $r = 4$, $u = 9$, $z = -3$, $w = 4$.

(ii) $ax + by = c \quad$ and $\quad dx + ey = f$.

4. **(i)** $\mathbf{A} + \mathbf{B}$

$$= \begin{bmatrix} 2 & 1 & 3 \\ 4 & 3 & 2 \\ 1 & 5 & 6 \end{bmatrix} + \begin{bmatrix} -1 & 4 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 2-1 & 1+4 & 3+0 \\ 4+1 & 3+1 & 2+0 \\ 1+0 & 5+1 & 6+1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 5 & 3 \\ 5 & 4 & 2 \\ 1 & 6 & 7 \end{bmatrix}.$$

(ii) Not possible—as **A** and **C** have different sizes.

(iii) Not possible—as **C** and **x** have different sizes.

(iv) $$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \left(\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ -3 \end{bmatrix} \right) + \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix}$$
$$= \begin{bmatrix} 3+a \\ 1+b \end{bmatrix}.$$

5.

$$\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ -3 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

implies the equations

$$2 + 1 = a \qquad \text{and} \qquad 4 - 3 = b.$$

6. **(i)** **C** must be the same size as **A** and **B**, so **C** will be an $m \times n$ matrix.

(ii) c_{ij}, the element in the ith row and jth column of **C**, is formed by adding the equivalent elements in **A** and **B**. So

$$c_{ij} = a_{ij} + b_{ij}.$$

7.

$$-\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 4 & -1 \end{bmatrix}.$$

In this case the zero matrix will be 2×2—i.e. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

8.

$$\mathbf{A} - \mathbf{B} = \begin{bmatrix} 3 & 2 \\ 1 & 2 \\ 5 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 3 \\ -1 & 1 \\ 2 & -3 \end{bmatrix}$$
$$= \begin{bmatrix} 3-0 & 2-3 \\ 1-(-1) & 2-1 \\ 5-2 & 1-(-3) \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 2 & 1 \\ 3 & 4 \end{bmatrix}.$$

9. **(i)** $$k\mathbf{A} = 5\begin{bmatrix} 2 & 1 \\ 3 & 2 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 5 \\ 15 & 10 \\ 20 & 5 \end{bmatrix}.$$

(ii) $$k(\mathbf{A} + m\mathbf{B}) = 5\mathbf{A} + 10\mathbf{B} = \begin{bmatrix} 10 & 5 \\ 15 & 10 \\ 20 & 5 \end{bmatrix} + \begin{bmatrix} -10 & 0 \\ 10 & 0 \\ 10 & -10 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 5 \\ 25 & 10 \\ 30 & -5 \end{bmatrix}.$$

(iii) $$k\mathbf{A} - m\mathbf{B} = \begin{bmatrix} 10 & 5 \\ 15 & 10 \\ 20 & 5 \end{bmatrix} - \begin{bmatrix} -2 & 0 \\ 2 & 0 \\ 2 & -2 \end{bmatrix}$$
$$= \begin{bmatrix} 12 & 5 \\ 13 & 10 \\ 18 & 7 \end{bmatrix}.$$

(iv) $$m\mathbf{A} + m\mathbf{B} + k\mathbf{A} - k\mathbf{B} = (m+k)\mathbf{A} + (m-k)\mathbf{B}$$
$$= 7\begin{bmatrix} 2 & 1 \\ 3 & 2 \\ 4 & 1 \end{bmatrix} + (-3)\begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 1 & -1 \end{bmatrix}$$
$$= \begin{bmatrix} 17 & 7 \\ 18 & 14 \\ 25 & 10 \end{bmatrix}.$$

10. **(i)** $$2(\mathbf{A} + 3\mathbf{B}) - 4(\mathbf{A} - \mathbf{B}) = 2\mathbf{A} + 6\mathbf{B} - 4\mathbf{A} + 4\mathbf{B}$$
$$= 10\mathbf{B} - 2\mathbf{A}.$$

(ii) $$\mathbf{X} + 3(\mathbf{A} + \mathbf{X}) = 2\mathbf{X} - \mathbf{A}$$
$$4\mathbf{X} + 3\mathbf{A} = 2\mathbf{X} - \mathbf{A}$$
$$2\mathbf{X} = -4\mathbf{A}$$
$$\mathbf{X} = -2\mathbf{A}.$$

11. **(i)** (a) $\mathbf{A} + \mathbf{B} = \begin{bmatrix} 2 & 4 \\ 7 & -1 \end{bmatrix} + \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 8 & 3 \\ 9 & 2 \end{bmatrix}$.

(b) $\mathbf{B} - \mathbf{C} = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & -3 \\ 2 & 2 \end{bmatrix}$.

(c) Not possible.

(d) $\mathbf{x} + 2\mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} x_1 + 4 \\ x_2 + 6 \end{bmatrix}$.

(e) $$\mathbf{A} - 2\mathbf{B} + \mathbf{C} = \begin{bmatrix} 2 & 4 \\ 7 & -1 \end{bmatrix} - \begin{bmatrix} 12 & -2 \\ 4 & 6 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} -9 & 8 \\ 3 & -6 \end{bmatrix}.$$

(ii) $2(\mathbf{P} + 2\mathbf{A}) + \mathbf{B} = 3\mathbf{C}$;

$$2\mathbf{P} + 4\mathbf{A} + \mathbf{B} = 3\mathbf{C};$$

$$\mathbf{P} = \tfrac{1}{2}(3\mathbf{C} - 4\mathbf{A} - \mathbf{B})$$

$$= \tfrac{1}{2}\left\{\begin{bmatrix} 3 & 6 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} 8 & 16 \\ 28 & -4 \end{bmatrix} - \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\right\}$$

$$= \tfrac{1}{2}\begin{bmatrix} -11 & -9 \\ -30 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{11}{2} & -\frac{9}{2} \\ -15 & 2 \end{bmatrix}.$$

Solutions to the exercises in Section 2

1. Using Tables 1 and 2:

(i) John's requirements in shop B cost

$$(6 \times 17) + (3 \times 32) + (2 \times 33) = 264\text{p}.$$

(ii) Jane's requirements in shop A cost

$$(10 \times 18) + (2 \times 30) + (2 \times 35) = 310\text{p}.$$

(iii) Jane's requirements in shop B cost

$$(10 \times 17) + (2 \times 32) + (2 \times 33) = 300\text{p}.$$

(iv) Joyce's requirements in shop A cost

$$(7 \times 18) + (2 \times 30) + (1 \times 35) = 221\text{p}.$$

(v) Jim's requirements in shop A cost

$$(15 \times 18) + (4 \times 30) + (3 \times 35) = 495\text{p}.$$

(vi) Jim's requirements in shop B cost

$$(15 \times 17) + (4 \times 32) + (3 \times 33) = 482\text{p}.$$

The completed table is in the main text, following Exercise 1.

2.

$$\mathbf{AB} = \begin{bmatrix} 4 & 7 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} (4 \times 3) + (7 \times -1) & (4 \times 1) + (7 \times 2) \\ (1 \times 3) + (2 \times -1) & (1 \times 1) + (2 \times 2) \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 18 \\ 1 & 5 \end{bmatrix}.$$

3. **AB** has the same number of rows as **A**, namely 3, and the same number of columns as **B**, namely 2, so **AB** is a 3×2 matrix.

4. **(i)** **B** will have to have 4 rows to form **AB** (no restrictions on how many columns), so **B** will be a $4 \times n$ matrix.

(ii) **B** will have to have 2 columns to form **BA** (no restrictions on the number of rows), so **B** will be an $m \times 2$ matrix.

5. **(i)** **AB**, **AC**, **BC**, **CA**. (**BA** and **CB** can't be formed.)

(ii) **AB** is a 1×2 matrix, **AC** is a 1×1 matrix, **BC** is a 2×1 matrix and **CA** is a 2×2 matrix.

6. Possible matrix products are **BC**, **BA**, **CB**.
(**AC**, **CA** and **AB** are not possible.)

$$\mathbf{BC} = [1 \quad -2]\begin{bmatrix} 3 \\ -1 \end{bmatrix} = [5].$$

$$\mathbf{BA} = [1 \quad -2]\begin{bmatrix} 2 & 1 & 0 \\ 3 & 2 & -1 \end{bmatrix} = [-4 \quad -3 \quad 2].$$

$$\mathbf{CB} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}[1 \quad -2] = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix}.$$

7. **(i)** $\mathbf{AB} = \begin{bmatrix} 2 & 0 & 1 \\ -1 & 2 & 1 \\ 3 & 4 & -2 \end{bmatrix}\begin{bmatrix} 1 & 2 & -1 \\ -1 & 2 & 3 \\ 3 & 4 & 1 \end{bmatrix}$

$$= \begin{bmatrix} 5 & 8 & -1 \\ 0 & 6 & 8 \\ -7 & 6 & 7 \end{bmatrix}.$$

(ii) $\mathbf{BA} = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 2 & 3 \\ 3 & 4 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 & 1 \\ -1 & 2 & 1 \\ 3 & 4 & -2 \end{bmatrix}$

$$= \begin{bmatrix} -3 & 0 & 5 \\ 5 & 16 & -5 \\ 5 & 12 & 5 \end{bmatrix}.$$

Note that **AB** and **BA** are not the same.

8. It is sufficient in these proofs to show that the ijth element in the left-hand matrix is the same as the equivalent one in the right-hand matrix. We will call such an element a 'typical' element.

(i) A typical element of $\mathbf{A}(\mathbf{B} + \mathbf{C})$ is obtained by taking the ith row of **A**,

$$[a_{i1} \quad a_{i2} \quad \ldots \quad a_{in}],$$

and combining it with the jth column of $\mathbf{B} + \mathbf{C}$,

$$\begin{bmatrix} b_{1j} + c_{1j} \\ b_{2j} + c_{2j} \\ \vdots \\ b_{nj} + c_{nj} \end{bmatrix}, \text{ to obtain}$$

$$a_{i1}(b_{1j} + c_{1j}) + a_{i2}(b_{2j} + c_{2j}) + \cdots + a_{in}(b_{nj} + c_{nj})$$

$$= \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$$

$$= \left(\sum_{k=1}^{n} a_{ik}b_{kj}\right) + \left(\sum_{k=1}^{n} a_{ik}c_{kj}\right).$$

As these expressions are the typical elements of **AB** and **AC**,

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}.$$

(ii) A typical element of $(\mathbf{B} + \mathbf{C})\mathbf{A}$ is formed by taking the ith row of $\mathbf{B} + \mathbf{C}$,

$$[b_{i1} + c_{i1} \quad b_{i2} + c_{i2} \quad \ldots \quad b_{in} + c_{in}],$$

and combining it with the jth column of **A**,

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}, \text{ to obtain}$$

$$(b_{i1} + c_{i1})a_{1j} + (b_{i2} + c_{i2})a_{2j} + \cdots + (b_{in} + c_{in})a_{nj}$$

$$= \sum_{k=1}^{n} (b_{ik} + c_{ik})a_{kj}$$

$$= \left(\sum_{k=1}^{n} b_{ik}a_{kj}\right) + \left(\sum_{k=1}^{n} c_{ik}a_{kj}\right).$$

As these are typical elements of **BA** and **CA**,

$$(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}.$$

(iii) This is the worst!

To find a typical element of $\mathbf{A}(\mathbf{BC})$, we combine the ith row of **A**,

$$[a_{i1} \quad a_{i2} \quad \ldots \quad a_{in}],$$

with the jth column of **BC**,

$$\begin{bmatrix} \sum_{k=1}^{n} b_{1k}c_{kj} \\ \sum_{k=1}^{n} b_{2k}c_{kj} \\ \vdots \\ \sum_{k=1}^{n} b_{nk}c_{kj} \end{bmatrix},$$

giving us the elements

$$\begin{aligned} a_{i1}\sum_{k=1}^{n} b_{1k}c_{kj} + a_{i2}\sum_{k=1}^{n} b_{2k}c_{kj} + \cdots + a_{in}\sum_{k=1}^{n} b_{nk}c_{kj} &= \sum_{q=1}^{n} a_{iq}\left(\sum_{k=1}^{n} b_{qk}c_{kj}\right) \\ &= \sum_{q=1}^{n}\sum_{k=1}^{n} a_{iq}b_{qk}c_{kj} \\ &= \sum_{k=1}^{n}\sum_{q=1}^{n} a_{iq}b_{qk}c_{kj} \\ &= \sum_{k=1}^{n}\left(\sum_{q=1}^{n} a_{iq}b_{qk}\right)c_{kj}. \end{aligned}$$

And this, believe it or not, is a typical element of $(\mathbf{AB})\mathbf{C}$.

9.

$$\mathbf{Aq} = \begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}.$$

10.

$$\begin{aligned} 3x_1 + x_2 - 2x_3 &= 2 \\ x_1 + 2x_2 + x_3 &= 1 \\ 4x_1 + x_2 \quad\quad &= -1 \end{aligned}$$

11.

$$\underset{\mathbf{A}}{\begin{bmatrix} 2 & -3 & 1 \\ 7 & 1 & -1 \\ 9 & 3 & 5 \end{bmatrix}}\underset{\mathbf{x}}{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}} \underset{=}{=} \underset{\mathbf{b}}{\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}}.$$

12.

$$\begin{aligned} \mathbf{r}_4 &= \mathbf{C}\mathbf{r}_3 \\ &= \mathbf{C}(\mathbf{B}\mathbf{r}_2) = (\mathbf{CB})\mathbf{r}_2 \\ &= \mathbf{CB}(\mathbf{A}\mathbf{r}_1) = (\mathbf{CBA})\mathbf{r}_1 \\ &= \mathbf{Q}\mathbf{r}_1 \quad \text{where } \mathbf{Q} = \mathbf{CBA}. \end{aligned}$$

$$\begin{aligned} \mathbf{Q} &= \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 0 & -4 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -3 & -5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -13 & 5 \\ 5 & -1 \end{bmatrix}. \end{aligned}$$

13.

$$\mathbf{A}^T = \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 7 & 2 \end{bmatrix} \text{ and } \mathbf{B}^T = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 3 & 7 \end{bmatrix}.$$

14.

$$\mathbf{AB} = \begin{bmatrix} 2 & 3 & 7 \\ 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 3 \\ -1 & 7 \end{bmatrix} = \begin{bmatrix} -3 & 60 \\ 0 & 15 \end{bmatrix},$$

so

$$(\mathbf{AB})^T = \begin{bmatrix} -3 & 0 \\ 60 & 15 \end{bmatrix}.$$

$$\mathbf{B}^T\mathbf{A}^T = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 3 & 7 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 7 & 2 \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 60 & 15 \end{bmatrix},$$

so for these matrices

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T.$$

(It is worth noting that $\mathbf{A}^T\mathbf{B}^T$ would be a 3×3 matrix, so there would be no hope that $(\mathbf{AB})^T$ equalled $\mathbf{A}^T\mathbf{B}^T$.)

15. (i) $\mathbf{Ab} = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}.$

(ii) $$\mathbf{AC} = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 4 \\ 1 & 2 & 3 \end{bmatrix}.$$

(iii) **AD** can't be formed.

(iv) **bC** can't be formed.

(v) $\mathbf{b}^T\mathbf{C} = [3 \quad 1]\begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} = [-3 \quad 1 \quad 5].$

(vi) $$\mathbf{DC}^T = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 1 \\ 2 & 3 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 8 \\ -3 & 7 \\ -1 & 5 \end{bmatrix}.$$

16. From Exercise 7,

$$\mathbf{AB} = \begin{bmatrix} 5 & 8 & -1 \\ 0 & 6 & 8 \\ -7 & 6 & 7 \end{bmatrix} \text{ and } \mathbf{BA} = \begin{bmatrix} -3 & 0 & 5 \\ 5 & 16 & -5 \\ 5 & 12 & 5 \end{bmatrix}.$$

Also,

$$\mathbf{AC} = \begin{bmatrix} 2 & 0 & 1 \\ -1 & 2 & 1 \\ 3 & 4 & -2 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 4 & 3 \\ 5 & 8 & -7 \end{bmatrix},$$

$$\mathbf{CA} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & 0 & 1 \\ -1 & 2 & 1 \\ 3 & 4 & -2 \end{bmatrix} = \begin{bmatrix} -1 & -4 & 3 \\ -2 & 4 & 2 \\ 4 & 8 & -5 \end{bmatrix},$$

$$\mathbf{BC} = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 2 & 3 \\ 3 & 4 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 4 & -3 \\ -4 & 4 & 7 \\ 2 & 8 & -1 \end{bmatrix}.$$

Thus:

(i)

$$(\mathbf{AB})\mathbf{C} = \begin{bmatrix} 5 & 8 & -1 \\ 0 & 6 & 8 \\ -7 & 6 & 7 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 16 & -7 \\ -8 & 12 & 16 \\ -14 & 12 & 21 \end{bmatrix},$$

$$\mathbf{A}(\mathbf{BC}) = \begin{bmatrix} 2 & 0 & 1 \\ -1 & 2 & 1 \\ 3 & 4 & -2 \end{bmatrix}\begin{bmatrix} 2 & 4 & -3 \\ -4 & 4 & 7 \\ 2 & 8 & -1 \end{bmatrix} = \begin{bmatrix} 6 & 16 & -7 \\ -8 & 12 & 16 \\ -14 & 12 & 21 \end{bmatrix},$$

so for the given matrices,

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}).$$

(ii)

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \begin{bmatrix} 2 & 0 & 1 \\ -1 & 2 & 1 \\ 3 & 4 & -2 \end{bmatrix}\begin{bmatrix} 2 & 2 & -2 \\ -1 & 4 & 3 \\ 2 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 6 & 8 & -1 \\ -2 & 10 & 11 \\ -2 & 14 & 0 \end{bmatrix},$$

$$\mathbf{AB}+\mathbf{AC} = \begin{bmatrix} 5 & 8 & -1 \\ 0 & 6 & 8 \\ -7 & 6 & 7 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ -2 & 4 & 3 \\ 5 & 8 & -7 \end{bmatrix} = \begin{bmatrix} 6 & 8 & -1 \\ -2 & 10 & 11 \\ -2 & 14 & 0 \end{bmatrix},$$

so for these matrices,

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}.$$

(iii)

$$(\mathbf{B}+\mathbf{C})\mathbf{A} = \begin{bmatrix} 2 & 2 & -2 \\ -1 & 4 & 3 \\ 2 & 4 & 3 \end{bmatrix}\begin{bmatrix} 2 & 0 & 1 \\ -1 & 2 & 1 \\ 3 & 4 & -2 \end{bmatrix} = \begin{bmatrix} -4 & -4 & 8 \\ 3 & 20 & -3 \\ 9 & 20 & 0 \end{bmatrix},$$

$$\mathbf{BA}+\mathbf{CA} = \begin{bmatrix} -3 & 0 & 5 \\ 5 & 16 & -5 \\ 5 & 12 & 5 \end{bmatrix} + \begin{bmatrix} -1 & -4 & 3 \\ -2 & 4 & 2 \\ 4 & 8 & -5 \end{bmatrix} = \begin{bmatrix} -4 & -4 & 8 \\ 3 & 20 & -3 \\ 9 & 20 & 0 \end{bmatrix},$$

so for these matrices,

$$(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA}+\mathbf{CA}.$$

17. (i)

$$\mathbf{AB} = \begin{bmatrix} 2 & 3 & 7 \\ 1 & 0 & -2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 3 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} 39 & 18 \\ -8 & -1 \end{bmatrix}$$

and

$$\mathbf{AC} = \begin{bmatrix} 2 & 3 & 7 \\ 1 & 0 & -2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 4 & -1 \end{bmatrix} = \begin{bmatrix} 36 & 2 \\ -7 & 2 \end{bmatrix},$$

so

$$\mathbf{AB}+\mathbf{AC} = \begin{bmatrix} 39 & 18 \\ -8 & -1 \end{bmatrix} + \begin{bmatrix} 36 & 2 \\ -7 & 2 \end{bmatrix} = \begin{bmatrix} 75 & 20 \\ -15 & 1 \end{bmatrix}.$$

Now,

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \begin{bmatrix} 2 & 3 & 7 \\ 1 & 0 & -2 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 2 & 6 \\ 9 & 0 \end{bmatrix} = \begin{bmatrix} 75 & 20 \\ -15 & 1 \end{bmatrix},$$

so the left-handed distributive law is true for these matrices.

(ii)

$$\mathbf{BA} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \\ 5 & 1 \end{bmatrix}\begin{bmatrix} 2 & 3 & 7 \\ 1 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 5 & 6 & 12 \\ 3 & 0 & -6 \\ 11 & 15 & 33 \end{bmatrix}$$

and

$$\mathbf{CA} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 4 & -1 \end{bmatrix}\begin{bmatrix} 2 & 3 & 7 \\ 1 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 7 \\ 7 & 6 & 8 \\ 7 & 12 & 30 \end{bmatrix},$$

so

$$\mathbf{BA}+\mathbf{CA} = \begin{bmatrix} 5 & 6 & 12 \\ 3 & 0 & -6 \\ 11 & 15 & 33 \end{bmatrix} + \begin{bmatrix} 2 & 3 & 7 \\ 7 & 6 & 8 \\ 7 & 12 & 30 \end{bmatrix} = \begin{bmatrix} 7 & 9 & 19 \\ 10 & 6 & 2 \\ 18 & 27 & 63 \end{bmatrix}.$$

Now,

$$(\mathbf{B}+\mathbf{C})\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 2 & 6 \\ 9 & 0 \end{bmatrix}\begin{bmatrix} 2 & 3 & 7 \\ 1 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 7 & 9 & 19 \\ 10 & 6 & 2 \\ 18 & 27 & 63 \end{bmatrix},$$

so the right-handed distributive law is true for these matrices.

(iii) The associative law does not apply, as **BC** can't be formed (the number of columns of **B** is not equal to the number of rows of **C**).

18.

$$\begin{bmatrix} 2 & 3 & 4 \\ 1 & -2 & 5 \\ 0 & 1 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 10 \end{bmatrix}.$$

19.

$$\begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 5 & 3 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 8 & -1 \\ 19 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}.$$

Solutions to the exercises in Section 3

1.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \text{ so } u = 1, v = 1.$$

2.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos 30^\circ & \sin 30^\circ \\ -\sin 30^\circ & \cos 30^\circ \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \sqrt{3}+\frac{3}{2} \\ -1 + \frac{3\sqrt{3}}{2} \end{bmatrix},$$

so

$$x' = \sqrt{3} + \frac{3}{2} \simeq 3.23,$$

$$y' = -1 + \frac{3\sqrt{3}}{2} \simeq 1.60.$$

3.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos 45^\circ & \sin 45^\circ \\ -\sin 45^\circ & \cos 45^\circ \end{bmatrix}\begin{bmatrix} 3-2 \\ 4-(-5) \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 1 \\ 9 \end{bmatrix} = \begin{bmatrix} \frac{10}{\sqrt{2}} \\ \frac{8}{\sqrt{2}} \end{bmatrix},$$

so

$$x' = \frac{10}{\sqrt{2}} \simeq 7.07,$$
$$y' = \frac{8}{\sqrt{2}} \simeq 5.66.$$

4.

$$\mathbf{M}_\gamma\mathbf{M}_\beta\mathbf{M}_\alpha = \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & \sin\beta \\ 0 & -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\gamma & \sin\gamma\cos\beta & \sin\gamma\sin\beta \\ -\sin\gamma & \cos\gamma\cos\beta & \cos\gamma\sin\beta \\ 0 & -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\gamma\cos\alpha - \sin\gamma\cos\beta\sin\alpha & \cos\gamma\sin\alpha + \sin\gamma\cos\beta\cos\alpha & \sin\gamma\sin\beta \\ -\sin\gamma\cos\alpha - \cos\gamma\cos\beta\sin\alpha & -\sin\gamma\sin\alpha + \cos\gamma\cos\beta\cos\alpha & \cos\gamma\sin\beta \\ \sin\beta\sin\alpha & -\sin\beta\cos\alpha & \cos\beta \end{bmatrix}$$

(Alternatively, $\mathbf{M}_\gamma\mathbf{M}_\beta\mathbf{M}_\alpha$ can be evaluated as $\mathbf{M}_\gamma(\mathbf{M}_\beta\mathbf{M}_\alpha)$.) From Equation (6) on p. 30 we now obtain the equations

$$\begin{aligned} x' &= (\cos\gamma\cos\alpha - \sin\gamma\cos\beta\sin\alpha)(x-a) \\ &\quad + (\cos\gamma\sin\alpha + \sin\gamma\cos\beta\cos\alpha)(y-b) \\ &\quad + (\sin\gamma\sin\beta)(z-c), \end{aligned}$$

$$\begin{aligned} y' &= (-\sin\gamma\cos\alpha - \cos\gamma\cos\beta\sin\alpha)(x-a) \\ &\quad + (-\sin\gamma\sin\alpha + \cos\gamma\cos\beta\cos\alpha)(y-b) \\ &\quad + (\cos\gamma\sin\beta)(z-c), \\ z' &= (\sin\beta\sin\alpha)(x-a) - (\sin\beta\cos\alpha)(y-b) \\ &\quad + (\cos\beta)(z-c). \end{aligned}$$

Solutions to the exercises in Section 4

1.

$$\mathbf{MN} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$
$$= \begin{bmatrix} \cos^2\alpha + \sin^2\alpha & -\cos\alpha\sin\alpha + \sin\alpha\cos\alpha \\ -\sin\alpha\cos\alpha + \cos\alpha\sin\alpha & \sin^2\alpha + \cos^2\alpha \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I},$$

and

$$\mathbf{NM} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}\begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}$$
$$= \begin{bmatrix} \cos^2\alpha + \sin^2\alpha & \cos\alpha\sin\alpha - \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha - \cos\alpha\sin\alpha & \sin^2\alpha + \cos^2\alpha \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}.$$

So $\mathbf{MN} = \mathbf{NM} = \mathbf{I}$.

2.

$$\mathbf{AB} = \frac{1}{ad-bc}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$
$$= \frac{1}{ad-bc}\begin{bmatrix} ad-bc & -ab+ba \\ cd-dc & -cb+da \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I},$$

and

$$\mathbf{BA} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$
$$= \frac{1}{ad-bc}\begin{bmatrix} da-bc & db-bd \\ -ca+ac & -cb+ad \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}.$$

So $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$.

3. (i) If $\mathbf{A} = \begin{bmatrix} 2 & 7 \\ 1 & 6 \end{bmatrix}$, use the formula for $\mathbf{A}^{-1}$ on p. 32 with $a = 2$, $b = 7$, $c = 1$, $d = 6$. This gives

$$\mathbf{A}^{-1} = \frac{1}{5}\begin{bmatrix} 6 & -7 \\ -1 & 2 \end{bmatrix}.$$

(ii) If $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}$, then $a = 1$, $b = -1$, $c = -2$, and $d = 2$. Thus,

$$ad - bc = (1 \times 2) - (-1 \times -2) = 0,$$

so the inverse can't be formed.

(iii) If $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, then $a = 0$, $b = 1$, $c = 1$, $d = 0$, and

$$\mathbf{A}^{-1} = \frac{1}{0-1}\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

(Note that in this case, $\mathbf{A}$ is its own inverse).

4. (i)

$$\begin{array}{r} \\ \\ \\ \mathbf{R}_2 - \frac{1}{2}\mathbf{R}_1 \end{array}
\begin{array}{cc|cc} \multicolumn{2}{c|}{\mathbf{A}} & \multicolumn{2}{c}{\mathbf{I}} \\ 2 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 \\ 0 & \frac{3}{2} & -\frac{1}{2} & 1 \end{array}
\begin{array}{l} \\ \mathbf{R}_1 \\ \mathbf{R}_2 \\ \mathbf{R}_1 \\ \mathbf{R}_{2a} \end{array}$$

$$\begin{array}{r} \\ \mathbf{R}_{2a} \div \frac{3}{2} \\ \mathbf{R}_1 - \mathbf{R}_{2b} \\ \\ \mathbf{R}_{1a} \div 2 \\ \\ \\ \end{array}
\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & 1 & -\frac{1}{3} & \frac{2}{3} \\ 2 & 0 & \frac{4}{3} & -\frac{2}{3} \\ 0 & 1 & -\frac{1}{3} & \frac{2}{3} \\ 1 & 0 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & -\frac{1}{3} & \frac{2}{3} \\ \multicolumn{2}{c|}{\mathbf{I}} & \multicolumn{2}{c}{\mathbf{A}^{-1}} \end{array}
\begin{array}{l} \mathbf{R}_1 \\ \mathbf{R}_{2b} \\ \mathbf{R}_{1a} \\ \mathbf{R}_{2b} \\ \mathbf{R}_{1b} \\ \mathbf{R}_{2b} \\ \\ \end{array}$$

So $\mathbf{A}^{-1} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$.

Check either that $\begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \mathbf{I}$

or $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} = \mathbf{I}$.

You don't have to do both; remember that $\mathbf{A}$ and $\mathbf{A}^{-1}$ commute.

(ii)

$$\begin{array}{r} \\ \\ \\ \end{array}\begin{array}{ccc|ccc} & \mathbf{A} & & & \mathbf{I} & \end{array}$$

$$\begin{array}{r} \\ \\ \end{array}\left[\begin{array}{ccc|ccc}1&3&0&1&0&0\\0&1&3&0&1&0\\3&0&1&0&0&1\end{array}\right]\begin{array}{l}\mathbf{R}_1\\\mathbf{R}_2\\\mathbf{R}_3\end{array}$$

$$\begin{array}{r} \\ \\ \mathbf{R}_3-3\mathbf{R}_1\end{array}\left[\begin{array}{ccc|ccc}1&3&0&1&0&0\\0&1&3&0&1&0\\0&-9&1&-3&0&1\end{array}\right]\begin{array}{l}\mathbf{R}_1\\\mathbf{R}_2\\\mathbf{R}_{3a}\end{array}$$

$$\begin{array}{r} \\ \\ \mathbf{R}_{3a}+9\mathbf{R}_2\end{array}\left[\begin{array}{ccc|ccc}1&3&0&1&0&0\\0&1&3&0&1&0\\0&0&28&-3&9&1\end{array}\right]\begin{array}{l}\mathbf{R}_1\\\mathbf{R}_2\\\mathbf{R}_{3b}\end{array}$$

$$\begin{array}{r} \\ \\ \mathbf{R}_{3b}\div 28\end{array}\left[\begin{array}{ccc|ccc}1&3&0&1&0&0\\0&1&3&0&1&0\\0&0&1&\frac{-3}{28}&\frac{9}{28}&\frac{1}{28}\end{array}\right]\begin{array}{l}\mathbf{R}_1\\\mathbf{R}_2\\\mathbf{R}_{3c}\end{array}$$

$$\begin{array}{r} \\ \mathbf{R}_2-3\mathbf{R}_c\\ \end{array}\left[\begin{array}{ccc|ccc}1&3&0&1&0&0\\0&1&0&\frac{9}{28}&\frac{1}{28}&\frac{-3}{28}\\0&0&1&\frac{-3}{28}&\frac{9}{28}&\frac{1}{28}\end{array}\right]\begin{array}{l}\mathbf{R}_1\\\mathbf{R}_{2a}\\\mathbf{R}_{3c}\end{array}$$

$$\begin{array}{r}\mathbf{R}_1-3\mathbf{R}_{2a}\\ \\ \end{array}\left[\begin{array}{ccc|ccc}1&0&0&\frac{1}{28}&\frac{-3}{28}&\frac{9}{28}\\0&1&0&\frac{9}{28}&\frac{1}{28}&\frac{-3}{28}\\0&0&1&\frac{-3}{28}&\frac{9}{28}&\frac{1}{28}\end{array}\right]\begin{array}{l}\mathbf{R}_{1a}\\\mathbf{R}_{2a}\\\mathbf{R}_{3c}\end{array}$$

$$\begin{array}{ccc|ccc} & \mathbf{I} & & & \mathbf{A}^{-1} & \end{array}$$

For safety, you should again check that the $\mathbf{A}^{-1}$ you have found is correct by finding $\mathbf{AA}^{-1}$ or $\mathbf{A}^{-1}\mathbf{A}$, and showing it comes to $\mathbf{I}$.

5.

$$(\mathbf{A}+\mathbf{B})(\mathbf{A}-\mathbf{B}) = \mathbf{A}^2+\mathbf{BA}-\mathbf{AB}-\mathbf{B}^2.$$

(Stop here, as $\mathbf{AB} \neq \mathbf{BA}$ in general.)

6.

$$\begin{aligned}&(\mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{ABC})\\ &\quad= \mathbf{C}^{-1}\mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{BC} \text{ (using the associative law)}\\ &\quad= \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{IBC} \qquad (\text{as } \mathbf{A}^{-1}\mathbf{A}=\mathbf{I})\\ &\quad= \mathbf{C}^{-1}\mathbf{IC} \qquad (\text{as } \mathbf{B}^{-1}\mathbf{B}=\mathbf{I})\\ &\quad= \mathbf{I} \qquad (\text{as } \mathbf{C}^{-1}\mathbf{C}=\mathbf{I}),\end{aligned}$$

or

$$(\mathbf{ABC})^{-1} = (\mathbf{A}(\mathbf{BC}))^{-1} = (\mathbf{BC})^{-1}\mathbf{A}^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}.$$

7. (i) A.

As $\mathbf{AA}^{-1}=\mathbf{I}$, by definition of an inverse, $\mathbf{A}$ must be the inverse of $\mathbf{A}^{-1}$.

(ii) A.

Let $\mathbf{A}^T=\mathbf{B}$ and $(\mathbf{A}^T)^T=\mathbf{C}$, and look at a typical element a_{ij} of $\mathbf{A}$. This becomes the element a_{ji} of $\mathbf{B}$ (as rows become columns). This in turn becomes the element a_{ij} of $\mathbf{C}$.

Thus $(\mathbf{A}^T)^T=\mathbf{A}$.

8.

$$\begin{aligned}(\mathbf{I}+\mathbf{A})(\mathbf{I}-\mathbf{A}) &= \mathbf{I}^2+\mathbf{AI}-\mathbf{IA}-\mathbf{A}^2\\ &= \mathbf{I}^2+\mathbf{A}-\mathbf{A}-\mathbf{I} \qquad\qquad (\text{as } \mathbf{A}^2=\mathbf{I}).\\ &= \mathbf{0}.\end{aligned}$$

9. We know from Subsection 2.5 that $(\mathbf{AB})^T=\mathbf{B}^T\mathbf{A}^T$. Letting $\mathbf{B}=\mathbf{A}^{-1}$, we have

$$(\mathbf{AA}^{-1})^T = (\mathbf{A}^{-1})^T\mathbf{A}^T.$$

Similarly,

$$(\mathbf{A}^{-1}\mathbf{A})^T = \mathbf{A}^T(\mathbf{A}^{-1})^T.$$

But $(\mathbf{AA}^{-1})^T$ and $(\mathbf{A}^{-1}\mathbf{A})^T$ are both $\mathbf{I}^T=\mathbf{I}$. This shows that $(\mathbf{A}^{-1})^T$ is the inverse of $\mathbf{A}^T$, which is what we want.

10. (i)

$$\begin{aligned}(\mathbf{A}+\mathbf{I})(\mathbf{A}-\mathbf{I}) &= \mathbf{A}^2+\mathbf{IA}-\mathbf{AI}-\mathbf{I}^2\\ &= \mathbf{A}^2-\mathbf{I}\end{aligned}$$

and

$$\begin{aligned}(\mathbf{A}-\mathbf{I})(\mathbf{A}+\mathbf{I}) &= \mathbf{A}^2-\mathbf{IA}+\mathbf{AI}-\mathbf{I}^2\\ &= \mathbf{A}^2-\mathbf{I}\\ &= (\mathbf{A}+\mathbf{I})(\mathbf{A}-\mathbf{I}).\end{aligned}$$

(ii) From part (i) we have

$$(\mathbf{A}+\mathbf{I})(\mathbf{A}-\mathbf{I}) = (\mathbf{A}-\mathbf{I})(\mathbf{A}+\mathbf{I}).$$

Multiply both sides on the left, and also on the right, by $(\mathbf{A}+\mathbf{I})^{-1}$.

$$\begin{aligned}&(\mathbf{A}+\mathbf{I})^{-1}(\mathbf{A}+\mathbf{I})(\mathbf{A}-\mathbf{I})(\mathbf{A}+\mathbf{I})^{-1}\\ &\qquad= (\mathbf{A}+\mathbf{I})^{-1}(\mathbf{A}-\mathbf{I})(\mathbf{A}+\mathbf{I})(\mathbf{A}+\mathbf{I})^{-1};\end{aligned}$$

so

$$\mathbf{I}(\mathbf{A}-\mathbf{I})(\mathbf{A}+\mathbf{I})^{-1} = (\mathbf{A}+\mathbf{I})^{-1}(\mathbf{A}-\mathbf{I})\mathbf{I};$$

i.e. $(\mathbf{A}-\mathbf{I})(\mathbf{A}+\mathbf{I})^{-1} = (\mathbf{A}+\mathbf{I})^{-1}(\mathbf{A}-\mathbf{I})$.

11. (i)

$$\begin{array}{cc|cc} & \mathbf{A} & \mathbf{I} & \end{array}$$

$$\left[\begin{array}{cc|cc}3&4&1&0\\9&12&0&1\end{array}\right]\begin{array}{l}\mathbf{R}_1\\\mathbf{R}_2\end{array}$$

$$\begin{array}{r} \\ \mathbf{R}_2-3\mathbf{R}_1\end{array}\left[\begin{array}{cc|cc}3&4&1&0\\0&0&-3&1\end{array}\right]\begin{array}{l}\mathbf{R}_1\\\mathbf{R}_{2a}\end{array}$$

$\mathbf{A}$ has no inverse.

(ii)

$$\begin{array}{cc|cc} & \mathbf{A} & \mathbf{I} & \end{array}$$

$$\left[\begin{array}{cc|cc}1&-2&1&0\\3&-1&0&1\end{array}\right]\begin{array}{l}\mathbf{R}_1\\\mathbf{R}_2\end{array}$$

$$\begin{array}{r} \\ \mathbf{R}_2-3\mathbf{R}_1\end{array}\left[\begin{array}{cc|cc}1&-2&1&0\\0&5&-3&1\end{array}\right]\begin{array}{l}\mathbf{R}_1\\\mathbf{R}_{2a}\end{array}$$

$$\begin{array}{r} \\ \mathbf{R}_{2a}\div 5\end{array}\left[\begin{array}{cc|cc}1&-2&1&0\\0&1&-\frac{3}{5}&\frac{1}{5}\end{array}\right]\begin{array}{l}\mathbf{R}_1\\\mathbf{R}_{2b}\end{array}$$

$$\begin{array}{r}\mathbf{R}_1+2\mathbf{R}_{2b}\\ \end{array}\left[\begin{array}{cc|cc}1&0&-\frac{1}{5}&\frac{2}{5}\\0&1&-\frac{3}{5}&\frac{1}{5}\end{array}\right]\begin{array}{l}\mathbf{R}_{1a}\\\mathbf{R}_{2b}\end{array}$$

$$\begin{array}{cc|cc} & \mathbf{I} & \mathbf{A}^{-1} & \end{array}$$

Hence

$$\mathbf{A}^{-1} = \begin{bmatrix}-\frac{1}{5} & \frac{2}{5}\\ -\frac{3}{5} & \frac{1}{5}\end{bmatrix}.$$

(iii)

$$
\begin{array}{r|ccc|ccc|l}
 & & \mathbf{A} & & & \mathbf{I} & & \\
 & 1 & 0 & 1 & 1 & 0 & 0 & \mathbf{R}_1 \\
 & 1 & 1 & 0 & 0 & 1 & 0 & \mathbf{R}_2 \\
 & 0 & 1 & 1 & 0 & 0 & 1 & \mathbf{R}_3 \\
\hline
 & 1 & 0 & 1 & 1 & 0 & 0 & \mathbf{R}_1 \\
\mathbf{R}_2-\mathbf{R}_1 & 0 & 1 & -1 & -1 & 1 & 0 & \mathbf{R}_{2a} \\
 & 0 & 1 & 1 & 0 & 0 & 1 & \mathbf{R}_3 \\
\hline
 & 1 & 0 & 1 & 1 & 0 & 0 & \mathbf{R}_1 \\
 & 0 & 1 & -1 & -1 & 1 & 0 & \mathbf{R}_{2a} \\
\mathbf{R}_3-\mathbf{R}_{2a} & 0 & 0 & 2 & 1 & -1 & 1 & \mathbf{R}_{3a} \\
\hline
 & 1 & 0 & 1 & 1 & 0 & 0 & \mathbf{R}_1 \\
 & 0 & 1 & -1 & -1 & 1 & 0 & \mathbf{R}_{2a} \\
\mathbf{R}_{3a}\div 2 & 0 & 0 & 1 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \mathbf{R}_{3b} \\
\hline
\mathbf{R}_1-\mathbf{R}_{3b} & 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \mathbf{R}_{1a} \\
\mathbf{R}_{2a}+\mathbf{R}_{3b} & 0 & 1 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \mathbf{R}_{2b} \\
 & 0 & 0 & 1 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \mathbf{R}_{3b} \\
 & & \mathbf{I} & & & \mathbf{A}^{-1} & &
\end{array}
$$

Hence

$$\mathbf{A}^{-1} = \tfrac{1}{2}\begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}.$$

12. We are given that

$$\mathbf{B} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}.$$

Multiply both sides on the left by $\mathbf{X}$ and on the right by $\mathbf{X}^{-1}$.

$$\mathbf{X}\mathbf{B}\mathbf{X}^{-1} = \mathbf{X}\mathbf{X}^{-1}\mathbf{A}\mathbf{X}\mathbf{X}^{-1} = \mathbf{A}.$$

Thus, $\mathbf{A} = \mathbf{X}\mathbf{B}\mathbf{X}^{-1}$.

13. **(i)**

$$
\begin{aligned}
\mathbf{A}^T\mathbf{B}^T &= (\mathbf{B}\mathbf{A})^T && \text{(property of transpose)} \\
&= (\mathbf{A}\mathbf{B})^T && \text{(as } \mathbf{A} \text{ and } \mathbf{B} \text{ commute)} \\
&= \mathbf{B}^T\mathbf{A}^T && \text{(property of transpose).}
\end{aligned}
$$

So $\mathbf{A}^T$ and $\mathbf{B}^T$ commute.

(ii)

$$
\begin{aligned}
\mathbf{A}^{-1}\mathbf{B}^{-1} &= (\mathbf{B}\mathbf{A})^{-1} && \text{(property of inverse)} \\
&= (\mathbf{A}\mathbf{B})^{-1} && \text{(as } \mathbf{A} \text{ and } \mathbf{B} \text{ commute)} \\
&= \mathbf{B}^{-1}\mathbf{A}^{-1} && \text{(property of inverse).}
\end{aligned}
$$

So $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ commute.

(iii)

$$
\begin{aligned}
\mathbf{A}\mathbf{B}^{-1} &= (\mathbf{B}^{-1}\mathbf{B})\mathbf{A}\mathbf{B}^{-1} && \text{(multiplying on the left by } \mathbf{I}) \\
&= \mathbf{B}^{-1}(\mathbf{B}\mathbf{A})\mathbf{B}^{-1} && \text{(using the associative rule)} \\
&= \mathbf{B}^{-1}(\mathbf{A}\mathbf{B})\mathbf{B}^{-1} && \text{(as } \mathbf{A} \text{ and } \mathbf{B} \text{ commute)} \\
&= \mathbf{B}^{-1}\mathbf{A}(\mathbf{B}\mathbf{B}^{-1}) && \text{(using the associative rule)} \\
&= \mathbf{B}^{-1}\mathbf{A}.
\end{aligned}
$$

So $\mathbf{A}$ and $\mathbf{B}^{-1}$ commute.

Solutions to the exercises in Section 5

1. **(i)** $\begin{vmatrix} 3 & 1 \\ -4 & 5 \end{vmatrix} = (3 \times 5) - (1 \times -4)$

$$= 15 + 4 = 19.$$

(ii) $\begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = (1 \times 2) - (2 \times 1) = 0.$

2. Using Equation (1):

(i)

$$
\begin{vmatrix} 1 & 2 & 0 \\ -1 & 2 & 3 \\ 4 & 1 & 5 \end{vmatrix} = 1\begin{vmatrix} 2 & 3 \\ 1 & 5 \end{vmatrix} - 2\begin{vmatrix} -1 & 3 \\ 4 & 5 \end{vmatrix}
$$

$$
\begin{aligned}
&= (1 \times (10-3)) - (2 \times (-5-12)) \\
&= 7 + 34 = 41.
\end{aligned}
$$

(ii)

$$
\det \mathbf{A}^T = \begin{vmatrix} 1 & -1 & 4 \\ 2 & 2 & 1 \\ 0 & 3 & 5 \end{vmatrix}
$$

$$
\begin{aligned}
&= 1\begin{vmatrix} 2 & 1 \\ 3 & 5 \end{vmatrix} + 1\begin{vmatrix} 2 & 1 \\ 0 & 5 \end{vmatrix} + 4\begin{vmatrix} 2 & 2 \\ 0 & 3 \end{vmatrix} \\
&= (1 \times 7) + (1 \times 10) + (4 \times 6) \\
&= 7 + 10 + 24 = 41.
\end{aligned}
$$

3.

$$
\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}
$$

$$
\begin{aligned}
&= 1\begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2\begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3\begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} \\
&= (1 \times -3) - (2 \times -6) + (3 \times -3) \\
&= -3 + 12 - 9 = 0.
\end{aligned}
$$

4. Using Equation (3):

$$
\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & -1 & 4 \\ 1 & -1 & 1 \end{vmatrix}
$$

$$
\begin{aligned}
&= \mathbf{i}\begin{vmatrix} -1 & 4 \\ -1 & 1 \end{vmatrix} - \mathbf{j}\begin{vmatrix} 3 & 4 \\ 1 & 1 \end{vmatrix} + \mathbf{k}\begin{vmatrix} 3 & -1 \\ 1 & -1 \end{vmatrix} \\
&= 3\mathbf{i} + \mathbf{j} - 2\mathbf{k}.
\end{aligned}
$$

5. The triangle OAB is the same as the triangle ABO, so all we have to do is to substitute the coordinates of O for those of C in Equation (6):

Area of triangle ABO

$$
\begin{aligned}
&= \tfrac{1}{2} \text{ modulus of } \begin{vmatrix} 1 & 1 & 1 \\ a_1 & b_1 & 0 \\ a_2 & b_2 & 0 \end{vmatrix} \\
&= \tfrac{1}{2} \text{ modulus of } \left\{\begin{vmatrix} b_1 & 0 \\ b_2 & 0 \end{vmatrix} - \begin{vmatrix} a_1 & 0 \\ a_2 & 0 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}\right\} \\
&= \tfrac{1}{2} \text{ modulus of } \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.
\end{aligned}
$$

6. **(i)** Using Equation (9):

$$
V = \text{modulus of } \begin{vmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 3 \end{vmatrix}
$$

$$= |7| = 7.$$

(ii) By part (i), $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = 7$.

$$
\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = \begin{vmatrix} 0 & 1 & 3 \\ 1 & 0 & 1 \\ 1 & 2 & 0 \end{vmatrix} = -1\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} + 3\begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix}
$$

$$= (-1 \times -1) + (3 \times 2) = 7.$$

$$
\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \begin{vmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 1 & 0 & 1 \end{vmatrix} = 1\begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} - 2\begin{vmatrix} 0 & 3 \\ 1 & 1 \end{vmatrix}
$$

$$= (1 \times 1) - (2 \times -3) = 7.$$

Hence

$$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})=\mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})=\mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}).$$

7.

$$\begin{aligned}\mathbf{c}\cdot(\mathbf{b}\times\mathbf{a}) &= -\mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}) && \text{(property of cross product)}\\ &= -\Delta && \text{(using (11)).}\end{aligned}$$

8.

$$\begin{aligned}\mathbf{a}\cdot(\mathbf{b}\times\mathbf{a}) &= \mathbf{b}\cdot(\mathbf{a}\times\mathbf{a}) && \text{(using (11))}\\ &= 0 && \text{(property of cross product).}\end{aligned}$$

9. If row **3** is a linear combination of rows 1 and 2, then $\mathbf{c}=k_1\mathbf{a}+k_2\mathbf{b}$, and so

$$\begin{aligned}\Delta &= \mathbf{a}\cdot(\mathbf{b}\times(k_1\mathbf{a}+k_2\mathbf{b}))\\ &= \mathbf{a}\cdot(\mathbf{b}\times(k_1\mathbf{a}))+\mathbf{a}\cdot(\mathbf{b}\times(k_2\mathbf{b}))\\ &= k_1(\mathbf{a}\cdot(\mathbf{b}\times\mathbf{a}))+k_2(\mathbf{a}\cdot(\mathbf{b}\times\mathbf{b})) && \text{(from P3)}\\ &= 0 && \text{(from P2).}\end{aligned}$$

10. (i)

$$\begin{vmatrix}1 & 2 & 7\\ 3 & 4 & 1\\ -1 & 2 & 7\end{vmatrix} = 1\begin{vmatrix}4 & 1\\ 2 & 7\end{vmatrix} - 2\begin{vmatrix}3 & 1\\ -1 & 7\end{vmatrix} + 7\begin{vmatrix}3 & 4\\ -1 & 2\end{vmatrix}$$

$$= (1\times 26)-(2\times 22)+(7\times 10)$$

$$= 26-44+70=52.$$

(ii)

$$\begin{vmatrix}1 & 2 & 7\\ 3 & 4 & 1\\ -1 & 2 & 7\end{vmatrix} = \begin{vmatrix}1 & 2 & 7\\ 0 & -2 & -20\\ 0 & 4 & 14\end{vmatrix} = \begin{vmatrix}1 & 2 & 7\\ 0 & -2 & -20\\ 0 & 0 & -26\end{vmatrix}$$

$$= 1\times -2\times -26 = 52.$$

11.

$$\begin{vmatrix}0 & 2 & 3\\ 1 & 3 & 6\\ 2 & 2 & 4\end{vmatrix} = -\begin{vmatrix}1 & 3 & 6\\ 0 & 2 & 3\\ 2 & 2 & 4\end{vmatrix} \quad \begin{pmatrix}\text{Essential interchange!}\\ \text{Remember to use P1.}\end{pmatrix}$$

$$= -\begin{vmatrix}1 & 3 & 6\\ 0 & 2 & 3\\ 0 & -4 & -8\end{vmatrix} = -\begin{vmatrix}1 & 3 & 6\\ 0 & 2 & 3\\ 0 & 0 & -2\end{vmatrix}$$

$$= -(1\times 2\times -2) = 4.$$

12. (i) $\begin{vmatrix}2 & 1\\ -1 & 4\end{vmatrix} = 8+1=9.$

(ii)

$$\begin{vmatrix}1 & 2 & -1\\ -1 & 1 & 2\\ 2 & -1 & 1\end{vmatrix} = 1\begin{vmatrix}1 & 2\\ -1 & 1\end{vmatrix} - 2\begin{vmatrix}-1 & 2\\ 2 & 1\end{vmatrix} - 1\begin{vmatrix}-1 & 1\\ 2 & -1\end{vmatrix}$$

$$= (1\times 3)-(2\times -5)-(1\times -1)$$

$$= 3+10+1$$

$$= 14.$$

(iii)

$$\begin{vmatrix}1 & 1 & 2 & 3\\ 1 & 2 & 4 & -1\\ 2 & 4 & -1 & 1\\ 0 & 1 & 2 & -1\end{vmatrix} = \begin{vmatrix}1 & 1 & 2 & 3\\ 0 & 1 & 2 & -4\\ 0 & 2 & -5 & -5\\ 0 & 1 & 2 & -1\end{vmatrix} = \begin{vmatrix}1 & 1 & 2 & 3\\ 0 & 1 & 2 & -4\\ 0 & 0 & -9 & 3\\ 0 & 0 & 0 & 3\end{vmatrix}$$

$$= 1\times 1\times -9\times 3$$

$$= -27.$$

Solutions to the end of unit test

Section A

1. The answer is (b). Option (c) is wrong because it is a 2×3 matrix. Remember, rows are specified before columns!

2. The answer is (d). (Equate corresponding coefficients!)

3. The answer is (d). (Only matrices of the same size can be added.)

4. The answer is (b). (Add corresponding elements.)

5. The answer is (a). (Multiply each element by 2.)

6. The answer is (a). (It is the only 1×1 matrix among the options, for a start!)

7. The answer is (c), being the only matrix among the options with 3 rows.

8. The answer is (c).

9. The answer is (e). This can be seen from the fact that the equations can be rewritten as

$$\begin{aligned}3x_1+2x_2+x_3&=4\\ x_2-x_3&=0\\ x_1+2x_2-x_3&=0.\end{aligned}$$

10. The answer is (c), as can be seen by multiplying both sides of $\mathbf{Ax}=\mathbf{b}$ on the left by $\mathbf{A}^{-1}$.

Section B

1. (i)

$$\mathbf{X}+\mathbf{I}=\begin{bmatrix}3 & 1 & 3\\ 1 & 0 & 0\\ 0 & 2 & 3\end{bmatrix}+\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} = \begin{bmatrix}4 & 1 & 3\\ 1 & 1 & 0\\ 0 & 2 & 4\end{bmatrix}.$$

(ii)

$$\mathbf{X}^2+\mathbf{X}=\mathbf{X}(\mathbf{X}+\mathbf{I}) = \begin{bmatrix}3 & 1 & 3\\ 1 & 0 & 0\\ 0 & 2 & 3\end{bmatrix}\begin{bmatrix}4 & 1 & 3\\ 1 & 1 & 0\\ 0 & 2 & 4\end{bmatrix} = \begin{bmatrix}13 & 10 & 21\\ 4 & 1 & 3\\ 2 & 8 & 12\end{bmatrix}.$$

(iii)

$$\mathbf{X}^3+\mathbf{X}^2=\mathbf{X}(\mathbf{X}^2+\mathbf{X}) = \begin{bmatrix}3 & 1 & 3\\ 1 & 0 & 0\\ 0 & 2 & 3\end{bmatrix}\begin{bmatrix}13 & 10 & 21\\ 4 & 1 & 3\\ 2 & 8 & 12\end{bmatrix} = \begin{bmatrix}49 & 55 & 102\\ 13 & 10 & 21\\ 14 & 26 & 42\end{bmatrix}.$$

2.

$$\begin{aligned}\mathbf{A}(\mathbf{I}+\mathbf{A})^2 &= \mathbf{A}(\mathbf{I}+2\mathbf{A}+\mathbf{A}^2)\\ &= \mathbf{A}(\mathbf{I}+2\mathbf{A}) && (\text{as } \mathbf{A}^2 = \mathbf{0})\\ &= \mathbf{A}+2\mathbf{A}^2\\ &= \mathbf{A} && (\text{as } \mathbf{A}^2 = \mathbf{0}).\end{aligned}$$

3. (i)

$$(\mathbf{A}+\mathbf{B})-(\mathbf{A}+\mathbf{C}) = \mathbf{A}+\mathbf{B}-\mathbf{A}-\mathbf{C} = \mathbf{B}-\mathbf{C}.$$

(ii)

$$\begin{aligned}&(\mathbf{A}+\mathbf{B})^2 - \mathbf{A}(\mathbf{A}+3\mathbf{B}) - \mathbf{B}(\mathbf{A}-\mathbf{B})\\ &= \mathbf{A}^2+\mathbf{AB}+\mathbf{BA}+\mathbf{B}^2-\mathbf{A}^2-3\mathbf{AB}-\mathbf{BA}+\mathbf{B}^2\\ &= 2\mathbf{B}^2-2\mathbf{AB}\\ &= 2(\mathbf{B}^2-\mathbf{AB})\\ &= 2(\mathbf{B}-\mathbf{A})\mathbf{B}.\end{aligned}$$

4. (i)

$$\begin{array}{ccc|ccc} & \mathbf{L} & & & \mathbf{I} & \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0\\ x & 1 & 0 & 0 & 1 & 0\\ y & z & 1 & 0 & 0 & 1\end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0\\ 0 & 1 & 0 & -x & 1 & 0\\ 0 & z & 1 & -y & 0 & 1\end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0\\ 0 & 1 & 0 & -x & 1 & 0\\ 0 & 0 & 1 & -y+xz & -z & 1\end{array}\right]$$

$$\begin{array}{ccc|ccc} & \mathbf{I} & & & \mathbf{L}^{-1} & \end{array}$$

So the inverse is

$$\begin{bmatrix} 1 & 0 & 0\\ -x & 1 & 0\\ -y+xz & -z & 1\end{bmatrix}.$$

(ii)

$$\mathbf{Bx} = \mathbf{Ax}+\mathbf{c};$$
$$(\mathbf{B}-\mathbf{A})\mathbf{x} = \mathbf{c}.$$

Now

$$\mathbf{B}-\mathbf{A} = \begin{bmatrix} 1 & 0 & 0\\ 1 & 1 & 0\\ 1 & 1 & 1\end{bmatrix}.$$

At this stage we can *either* use Gaussian elimination (giving us $x_1 = 3$, $x_2 = 1$, $x_3 = 1$) *or* proceed as follows.

$$(\mathbf{B}-\mathbf{A})\mathbf{x} = \mathbf{c};$$
$$\mathbf{x} = (\mathbf{B}-\mathbf{A})^{-1}\mathbf{c},$$

and putting $x = y = z = 1$ in the result of part (i),

$$\mathbf{x} = \begin{bmatrix} 1 & 0 & 0\\ -1 & 1 & 0\\ 0 & -1 & 1\end{bmatrix}\begin{bmatrix}3\\4\\5\end{bmatrix} = \begin{bmatrix}3\\1\\1\end{bmatrix}.$$

5. (i)

$$\mathbf{A}^2-\mathbf{A}+\mathbf{I} = \mathbf{0};$$
$$\mathbf{A}-\mathbf{A}^2 = \mathbf{I};$$
$$\mathbf{A}(\mathbf{I}-\mathbf{A}) = \mathbf{I};$$

so $\mathbf{I}-\mathbf{A}$ is the inverse of $\mathbf{A}$.

(ii)

$$\begin{aligned}(\mathbf{R}^{-1}\mathbf{AR})^3 &= (\mathbf{R}^{-1}\mathbf{AR})(\mathbf{R}^{-1}\mathbf{AR})(\mathbf{R}^{-1}\mathbf{AR})\\ &= \mathbf{R}^{-1}\mathbf{A}(\mathbf{RR}^{-1})\mathbf{A}(\mathbf{RR}^{-1})\mathbf{AR} && (\text{associative rule})\\ &= \mathbf{R}^{-1}\mathbf{A}^3\mathbf{R} && (\text{as } \mathbf{RR}^{-1} = \mathbf{I}).\end{aligned}$$

The general result is

$$(\mathbf{R}^{-1}\mathbf{AR})^n = \mathbf{R}^{-1}\mathbf{A}^n\mathbf{R} \qquad \text{for any positive integer } n.$$

6. (i)

$$\mathbf{AA}^T = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & -\frac{1}{2}\end{bmatrix}\begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2}\\ -\frac{\sqrt{3}}{2} & -\frac{1}{2}\end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{4}+\frac{3}{4} & -\frac{\sqrt{3}}{4}+\frac{\sqrt{3}}{4}\\ -\frac{\sqrt{3}}{4}+\frac{\sqrt{3}}{4} & \frac{3}{4}+\frac{1}{4}\end{bmatrix} = \mathbf{I}.$$

(Note, the inverse of an orthogonal matrix is its transpose.)

(ii)

$\mathbf{A}^{-1} = \mathbf{A}^T$ by part (i)

$$= \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2}\\ -\frac{\sqrt{3}}{2} & -\frac{1}{2}\end{bmatrix}.$$

(iii)

$$\begin{aligned}(\mathbf{AB})(\mathbf{AB})^T &= (\mathbf{AB})(\mathbf{B}^T\mathbf{A}^T) && (\text{property of transpose})\\ &= \mathbf{A}(\mathbf{BB}^T)\mathbf{A}^T && (\text{associative rule})\\ &= \mathbf{AA}^T && (\text{because } \mathbf{B} \text{ is an orthogonal matrix})\\ &= \mathbf{I} && (\text{because } \mathbf{A} \text{ is an orthogonal matrix}).\end{aligned}$$

Thus $\mathbf{AB}$ is an orthogonal matrix.

7. (i) $\mathbf{x}^T$ is $1 \times n$ and $\mathbf{Ax}$ is $n \times 1$. Thus, $\mathbf{x}^T\mathbf{Ax}$ is 1×1. (In practice, 1×1 matrices are usually thought of just as numbers.)

(ii)

$$\begin{aligned}\mathbf{Q} &= [x_1 \quad x_2]\begin{bmatrix} a & b\\ c & d\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}\\ &= [x_1 \quad x_2]\begin{bmatrix} ax_1+bx_2\\ cx_1+dx_2\end{bmatrix}\\ &= [x_1(ax_1+bx_2)+x_2(cx_1+dx_2)]\\ &= [ax_1^2+(b+c)x_1x_2+dx_2^2].\end{aligned}$$

(iii) Using part (ii), we must have $a = 1$, $d = 1$, $b + c = 2$. There are many ways of choosing b and c; for example, $b = c = 1$, giving

$$\mathbf{B} = \begin{bmatrix} 1 & 1\\ 1 & 1\end{bmatrix}.$$

8. (i) $\mathbf{y}^T\mathbf{Ax}$ is a matrix with one element, so $(\mathbf{y}^T\mathbf{Ax})^T$ is the same one-element matrix, i.e.

$$(\mathbf{y}^T\mathbf{Ax})^T = \mathbf{y}^T\mathbf{Ax}.$$

(ii)

$$\begin{aligned}&(\mathbf{y}-\mathbf{Ax})^T(\mathbf{y}-\mathbf{Ax})\\ &= (\mathbf{y}^T-\mathbf{x}^T\mathbf{A}^T)(\mathbf{y}-\mathbf{Ax}) && (\text{property of transpose})\\ &= \mathbf{y}^T\mathbf{y}-\mathbf{y}^T\mathbf{Ax}-\mathbf{x}^T\mathbf{A}^T\mathbf{y}+\mathbf{x}^T\mathbf{A}^T\mathbf{Ax}\\ &= \mathbf{y}^T\mathbf{y}-\mathbf{y}^T\mathbf{Ax}-(\mathbf{y}^T\mathbf{Ax})^T+\mathbf{x}^T\mathbf{A}^T\mathbf{Ax}\\ & && (\text{property of transpose})\\ &= \mathbf{y}^T\mathbf{y}-\mathbf{y}^T\mathbf{Ax}-\mathbf{y}^T\mathbf{Ax}+\mathbf{x}^T\mathbf{A}^T\mathbf{Ax} && (\text{from part (i)})\\ &= \mathbf{y}^T\mathbf{y}-2\mathbf{y}^T\mathbf{Ax}+\mathbf{x}^T\mathbf{A}^T\mathbf{Ax}.\end{aligned}$$

9. (i) $\begin{vmatrix} 4 & 1\\ -2 & 8\end{vmatrix} = (4 \times 8) - (1 \times -2) = 34.$

(ii)
$$\begin{vmatrix} 1 & 0 & -3 \\ -1 & 1 & 0 \\ 2 & 3 & -1 \end{vmatrix} = 1\begin{vmatrix} 1 & 0 \\ 3 & -1 \end{vmatrix} - 3\begin{vmatrix} -1 & 1 \\ 2 & 3 \end{vmatrix}$$
$$= (1 \times -1) - (3 \times -5) = 14.$$

(iii)
$$\begin{vmatrix} 1 & 3 & 0 & 4 \\ 0 & -1 & 2 & 1 \\ 3 & 0 & 1 & -1 \\ -2 & 3 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 0 & 4 \\ 0 & -1 & 2 & 1 \\ 0 & -9 & 1 & -13 \\ 0 & 9 & 1 & 8 \end{vmatrix}$$
$$= \begin{vmatrix} 1 & 3 & 0 & 4 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & -17 & -22 \\ 0 & 0 & 19 & 17 \end{vmatrix}$$
$$= \begin{vmatrix} 1 & 3 & 0 & 4 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & -17 & -22 \\ 0 & 0 & 0 & -\frac{129}{17} \end{vmatrix}$$
$$= 1 \times -1 \times -17 \times -\tfrac{129}{17}$$
$$= -129.$$

The Open University
0 335 14049 1